

1-1-2015

# Essays in Asset Pricing

Darien Huang
*University of Pennsylvania*, darien.huang@gmail.com

# Essays in Asset Pricing

**Abstract**
In the first chapter ``Gold, Platinum, and Expected Stock Returns'', I show that the ratio of gold to platinum prices (GP) reveals variation in risk and proxies for an important economic state variable. GP predicts future stock returns in the time-series and explains variation in average stock returns in the cross-section. GP outperforms existing predictors and similar patterns are found in international markets. GP is persistent and significantly correlated with option-implied tail risk measures. An equilibrium model featuring recursive preferences, time-varying tail risk, and shocks to preferences for gold and platinum can account for the asset pricing dynamics of equity, gold, and platinum markets, and quantitatively explain the return predictability. In the second chapter ``Risk Adjustment and the Temporal Resolution of Uncertainty: Evidence from Options Markets'', we examine risk-neutral probabilities, which are observable from option prices and combine objective probabilities and risk adjustments across economic states. We consider a recursive-utility framework to separately identify objective probabilities and risk adjustments using only observed market prices. We find that a preference for early resolution of uncertainty is important in explaining the cross-section of risk-neutral and objective probabilities in the data. Failure to incorporate a preference for the timing of the resolution of uncertainty (e.g., expected utility models) can significantly overstate the implied probability of, and understate risk compensations for, adverse economic states. In the third chapter ``Volatility-of-Volatility Risk'', we show that time-varying volatility of volatility is a significant risk factor which affects the cross-section and time-series of index and VIX option returns, beyond volatility risk itself. Volatility and volatility-of-volatility movements are identified from index and VIX option prices, and correspond to the VIX and VVIX indices in the data. Delta-hedged returns for index and VIX options are negative on average, and more negative for strategies more exposed to volatility and volatility-of-volatility risks. In the time-series, volatility and volatility of volatility significantly predict delta-hedged returns with a negative sign. The evidence is consistent with a no-arbitrage model featuring time-varying volatility and volatility-of-volatility factors which are negatively priced by investors.

**Degree Type**
Dissertation

**Degree Name**
Doctor of Philosophy (PhD)

**Graduate Group**
Finance

**First Advisor**
Amir Yaron

**Subject Categories**
Economics | Finance and Financial Management

# ESSAYS IN ASSET PRICING

Darien Huang

A DISSERTATION

in

Finance

For the Graduate Group in Managerial Science and Applied Economics

Presented to the Faculties of the University of Pennsylvania

in

Partial Fulfillment of the Requirements for the

Degree of Doctor of Philosophy

2015

Supervisor of Dissertation

---

Amir Yaron, Robert Morris Professor of Banking and Finance

Graduate Group Chairperson

---

Eric Bradlow, Professor of Marketing, Statistics, and Education

Dissertation Committee

Amir Yaron, Robert Morris Professor of Banking and Finance

Franklin Allen, Nippon Life Professor of Finance and Economics

Ivan Shaliastovich, Assistant Professor of Finance

*Dedicated to my parents.*

ACKNOWLEDGEMENT

I would like to thank my advisors Franklin Allen, Ivan Shaliastovich, and Amir Yaron (Chair) for their help, support, and guidance. I also thank Andy Abel, Erik Gilje, Itay Goldstein, Mete Kilic, Nick Roussanov, Luke Taylor, Rob Stambaugh, and Jessica Wachter for helpful discussions.

ABSTRACT

ESSAYS IN ASSET PRICING

Darien Huang

Amir Yaron

In the first chapter "Gold, Platinum, and Expected Stock Returns", I show that the ratio of gold to platinum prices (GP) reveals variation in risk and proxies for an important economic state variable. GP predicts future stock returns in the time-series and explains variation in average stock returns in the cross-section. GP outperforms existing predictors and similar patterns are found in international markets. GP is persistent and significantly correlated with option-implied tail risk measures. An equilibrium model featuring recursive preferences, time-varying tail risk, and shocks to preferences for gold and platinum can account for the asset pricing dynamics of equity, gold, and platinum markets, and quantitatively explain the return predictability.

In the second chapter "Risk Adjustment and the Temporal Resolution of Uncertainty: Evidence from Options Markets", we examine risk-neutral probabilities, which are observable from option prices and combine objective probabilities and risk adjustments across economic states. We consider a recursive-utility framework to separately identify objective probabilities and risk adjustments using only observed market prices. We find that a preference for early resolution of uncertainty is important in explaining the cross-section of risk-neutral and objective probabilities in the data. Failure to incorporate a preference for the timing of the resolution of uncertainty (e.g., expected utility models) can significantly overstate the implied probability of, and understate risk compensations for, adverse economic states.

In the third chapter "Volatility-of-Volatility Risk", we show that time-varying volatility of volatility is a significant risk factor which affects the cross-section and time-series of index and VIX option returns, beyond volatility risk itself. Volatility and volatility-of-volatility

movements are identified from index and VIX option prices, and correspond to the VIX and VVIX indices in the data. Delta-hedged returns for index and VIX options are negative on average, and more negative for strategies more exposed to volatility and volatility-of-volatility risks. In the time-series, volatility and volatility of volatility significantly predict delta-hedged returns with a negative sign. The evidence is consistent with a no-arbitrage model featuring time-varying volatility and volatility-of-volatility factors which are negatively priced by investors.

TABLE OF CONTENTS

# LIST OF TABLES

# LIST OF ILLUSTRATIONS

# CHAPTER 1 : Gold, Platinum, and Expected Stock Returns

## 1.1. Introduction

> *"As gold's unquenchable beauty shines like the sun, people have turned to it to protect themselves against the darkness ahead."*
> — Bernstein (2012), *The Power of Gold: The History of an Obsession*

Gold is one of the most important assets in financial markets and the global economy. As the author Peter Bernstein summarizes above, gold is viewed as two things: it is a consumption good (mostly jewellery) and it is also seen as something valuable in times of severe distress. Platinum, on the other hand, is a precious metal with similar uses as gold in consumption. Therefore, the ratio of gold to platinum prices should be largely insulated from shocks to consumption and jewelry demand, and should instead reveal variation in risk and proxy for an important economic state variable. I investigate three questions in this paper.

First, I ask whether the ratio of gold to platinum prices (GP) predicts future stock returns in the time-series and explains variation in average stock returns in the cross-section. I show empirically that GP is a strong predictor of future stock returns. A one standard deviation increase in GP predicts a 6.4% increase in U.S. stock market excess returns over the following year. GP outperforms nearly all existing return predictors and is robust to various econometric inference concerns highlighted in the literature. Gold and platinum are actively traded around the world, and similar patterns of stock return predictability are found in international markets. GP risk is priced in the cross-section of stock returns and commands a negative market price of risk.

After discussing the main empirical results, examining the mechanism which drives the results leads to my second question: Is gold a hedge? More specifically, do gold prices go up in bad times?[1] The answer - contrary to conventional wisdom - is no. Figure 1.1 plots real gold (top panel) and platinum (bottom panel) prices alongside stock market valuations

[1] See e.g., Erb and Harvey (2013), Barro and Misra (2013).

and NBER recession indicators from 1975 - 2013.[2] We see in the data that gold prices fall in recessions, albeit by less than platinum prices. For example, in the 1981 - 1982 recession, real gold prices fell 32% peak to trough, and in the recent 2008 - 2009 financial crisis real gold prices fell 22%. Unlike index put options or VIX futures, gold futures would not have helped investors hedge downside risks during the crises. Not by coincidence, the real price of platinum fell by 39% and 59% over the same periods, respectively.

To the extent that investors view shocks to gold prices as short-lived, flight-to-liquidity phenomena, I find that this is not the case. Shocks to GP do not correlate with shocks to transient measures of liquidity risk such as the Pastor and Stambaugh (2003) factor, and instead have a much longer half-life. Furthermore, GP is significantly related to measures of economic tail risk including the slope of the implied volatility curve for S&P500 index options, and the Bakshi, Kapadia, and Madan (2003) model-free risk-neutral skewness.[3] These findings lead to my final question, which is whether an extension of the time-varying disaster risk model (Wachter (2013)), which features recursive preferences and stochastic disaster probabilities, can quantitatively explain the time variation and return predictability of GP while simultaneously accounting for the asset pricing dynamics of equity, gold, and platinum markets, without any additional risk factors. The model is motivated by the fact that, under no arbitrage, investors are indifferent between buying gold or leasing gold in perpetuity (Barro and Misra (2013)).

I adopt a three-good model where agents derive utility from nondurable consumption as well as service flows from gold and platinum, which are non-depreciating durable goods with negligible outlays relative to nondurable consumption. In normal times, service flows from gold and platinum (which can be thought of as jewellery) complement nondurable

[2] I focus exclusively on the post-gold standard era, where gold prices vary freely by a market mechanism. While the "Nixon shock" of 1971 temporarily suspended convertibility of U.S. dollars into gold at $35 per oz, a new peg was later put in place at $38 per troy oz, followed by $42.22. Gold convertibility was only completely abolished by November 1973 (Lannoye (2011)). Executive Order 6102, put in place by President Franklin Roosevelt in 1933, banned gold trading within the United States. This act was repealed by President Gerald Ford in 1974 and took effect on December 31st, 1974. See Public Law 93-373.

[3] Tail risk is also known as jump risk or disaster risk depending on the literature.

consumption and are highly procyclical. However, when the probability of a consumption disaster is high, agents display an increased preference for gold relative to platinum. This is motivated by historical and institutional reasons, since gold is viewed as financial collateral and is formally recognized as such by the Basel Accords.[4]

The countercyclical benefits to physical ownership of gold and platinum are modeled in reduced-form using a pair of stochastic processes which are proportional to the probability of a consumption disaster; gold is calibrated to have greater countercyclical benefits than platinum, which is both consistent with the historical and institutional facts and also allows the model to rationalize the low gold lease rate and risk premium observed in the data.[5] In the model, GP is insulated from shocks to consumption since they affect gold and platinum prices equally. Increases in disaster probabilities raise risk premia, leading to higher discount rates and lower stock prices. Gold and platinum prices fall as well because of strong discount rate effects, although gold prices fall by less than platinum prices due to the higher countercyclical component of its service flow. As a result, GP is high when stock prices are low and the equity risk premium is high, giving GP the power to predict future stock market excess returns. The model quantitatively captures the key moments of gold and platinum returns, while remaining consistent with standard asset pricing moments such as the equity premium and risk-free rate. This is achieved without introducing additional state variables, which suggests that gold and platinum returns can largely be explained by the same risk factors affecting stocks and bonds.

Barro and Misra (2013) study gold returns in a Lucas (1978) endowment economy with rare consumption disasters. The authors match the low gold risk premium using a high elasticity of substitution between gold service flow and nondurable consumption. This assumption is problematic for two reasons. First, viewing gold as jewellery suggests a complementary

[4]See e.g., Basel I (1988), Basel II (2004), and Basel III (2012). Gold is also accepted as collateral by major derivatives exchanges and clearinghouses such as the CME and ICE Clear Europe, as well as large broker dealers such as JP Morgan.

[5]Platinum is not eligible collateral under the Basel Accords, central banks are not known to hold platinum reserves, and major financial institutions do not accept platinum as collateral.

rather than substitutable relationship (one cannot wear jewellery in place of consuming food, but jewellery is highly valued when food is plentiful). Second, optimality conditions reveal that the elasticity of substitution is inversely proportional to the degree of consumption leverage. Following analysis similar to Wachter (2013), I show that substitutability results in the counterfactual prediction that gold lease rates fall (gold prices rise) when disaster probabilities increase.

This paper contributes to the literature on return predictability by demonstrating that GP, a model-free measure available in real-time, is robust to and in most cases outperforms existing forecasting variables including equity valuation ratios (in various forms), the default spread, term spread, inflation, implied cost of capital, consumption-wealth ratio, and variance premium.[6] The predictive power of GP is stable both out-of-sample and over sub-samples, which alleviates concerns raised by studies such as Goyal and Welch (2008), who show that many predictors such as valuation ratios have low forecasting performance out-of-sample and unstable forecasting ability over sub-samples.

This paper extends the growing literature on gold and gold lease rates. To my knowledge, Barro and Misra (2013) is the only other paper to value gold in an equilibrium model. Fama and French (1988) analyze the behavior of metals prices over the business cycle based on the Brennan (1958) theory of storage. While base metals such as aluminum and copper behave as the theory of storage predicts, precious metals such as gold seem unresponsive; Fama and French hypothesize that this is due to low storage costs for precious metals.Tufano (1996) studies risk management practices in the gold mining industry. Schwartz (1997), Casassus and Collin-Dufresne (2005), and Le and Zhu (2013) study gold lease rates (known as "convenience yields" in the commodities literature) using dynamic term structure models.

[6]Valuation ratios include the price-dividend ratio, price-earnings ratio, and net payout yield. References include Campbell and Shiller (1988), Hodrick (1992), and Boudoukh, Michaely, Richardson, and Roberts (2007). References for predictability using business cycle variables include Lintner (1975), Campbell (1987), Keim and Stambaugh (1986), and Fama and French (1989). The implied cost of capital is studied by Pastor, Sinha, and Swaminathan (2008) and Li, Ng, and Swaminathan (2013). The consumption-wealth ratio (CAY) is from Lettau and Ludvigson (2001). References for variance premium predictability include Bollerslev, Tauchen, and Zhou (2009), Drechsler and Yaron (2011), and Drechsler (2013).

Erb and Harvey (2013) examine various theories regarding gold returns, including whether gold prices appreciate when stock prices fall. The authors find that many of the largest S&P500 declines were associated with falling gold prices.

Finally, this paper draws on the literature examining the impact of heavy-tailed shocks to economic state variables on asset prices. Examples from the option pricing literature include Bates (2000), Duffie, Pan, and Singleton (2000), Pan (2002), and Broadie, Chernov, and Johannes (2007). Jurek (2014) discusses the impact of crash risk on currency carry trade returns. Examples from the general equilibrium literature include the rare disasters framework (Rietz (1988), Barro (2006), Gabaix (2012), Gourio (2012), Wachter (2013), Nowotny (2011), Seo and Wachter (2014)), as well as extensions of the Bansal and Yaron (2004) long-run risks framework incorporating jumps in economic fundamentals (Eraker and Shaliastovich (2008), Bansal and Shaliastovich (2011), Benzoni, Collin-Dufresne, and Goldstein (2011), Drechsler and Yaron (2011)).

The paper proceeds as follows; data sources are discussed as the relevant sections are presented. Section 2 presents the empirical results on stock return predictability, cross-sectional evidence, and the relationship between GP and tail risk measures. Section 3 discusses key aspects of gold and platinum markets, focusing on sources of demand for each metal, the leasing markets, and return dynamics. Section 4 presents the model. Section 5 discusses the model calibration and simulation results. Section 6 concludes.

## 1.2. Empirical Results

### *1.2.1. Data Description*

Gold and platinum prices are the monthly average of daily fixing prices from the London Bullion Market Association (LBMA) and London Platinum and Palladium Market (LPPM), respectively, from 1975 to 2013.[7] Platinum fixing prices are available from April 1990; prior

[7] I use prices from the a.m. fixing, which is conducted at 9:45 a.m. GMT (for platinum) and 10:30 a.m. GMT (for gold).

to this, for platinum prices I use dealer prices from the U.S. Geological Survey.[8] The log GP ratio is calculated as the natural logarithm of the ratio of gold to platinum prices.[9] My measure of U.S. stock returns is the CRSP value-weighted index. The risk-free rate is the 1 month U.S. Treasury bill rate. I compare the performance of GP to various forecasting variables proposed in the literature.

- Price-Dividend Ratio ($\log PD$) is the log ratio of aggregate stock market price divided by the sum of the past twelve months of dividends. Dividends are computed from the difference between the CRSP value-weighted index return including and exluding dividends.

- Price-Earnings Ratio ($\log PE$) is the cyclically-adjusted log ratio of aggregate stock market price divided by past earnings, obtained from Robert Shiller's website.

- Net Payout Ratio ($\log PNY$) is the log ratio of total market capitalization divided by the sum of dividends, repurchases, and share issuance, as described in Boudoukh et al. (2007) and obtained from Michael Roberts's website. The series is available until December 2010.

- Implied Cost of Capital ($ICC$) is rate of return which solves the long horizon dividend discount model, constructed from I/B/E/S analyst earnings per share forecasts, as described in Li et al. (2013). The series starts from January 1977.

- Default Spread ($DFSP$) is the percentage difference in yield between Moody's Baa and Aaa rated corporate bonds and obtained from the Federal Reserve Bank of St. Louis (FRED) website.

- Term Spread ($TMSP$) is the percentage difference in yield between 10 year U.S. government bonds and 3 month U.S. Treasury bills and obtained from FRED.

[8] My results are nearly unchanged using platinum prices directly obtained from Platts, which is a large data vendor for metals markets.

[9] I use the terms "GP" and "log GP" interchangeably unless otherwise noted.

- Inflation ($INFL$) is the log growth rate of the *Consumer Price Index (All Urban Consumers: All Items)*, in percentages, from FRED.

- Consumption-Wealth Ratio ($CAY$) is the Lettau and Ludvigson (2001) measure of the consumption-wealth ratio, obtained from Martin Lettau's website. Monthly observations are computed by interpolating the quarterly observations. The series is available until March 2013.

- Variance Premium ($VRP$) is the difference between model-free implied variance computed from S&P500 option prices ($VIX^2$) and realized variance computed from 5-minute tick data over the past 30 days. The data for the VIX is obtained from the CBOE website, and the data for realized variance is from Hao Zhou's website. The series starts from January 1990.

Figure 1.2 plots the time-series of GP (solid line) along with the price-dividend ratio (dashed line). The average level of GP is below zero; gold trades at a 20% discount to platinum on average, consistent with platinum being a much scarcer metal. GP is strongly countercyclical and peaks during times of economic and financial distress including all NBER recessions between 1975 - 2013, as well as the October 1987 stock market crash, 1998 Russian default and LTCM crisis, and 2011 U.S. debt ceiling crisis. Table 1.1 presents summary statistics for all the predictors. With the exception of the variance premium and inflation, all other predictors are quite persistent. The AR(1) coefficient for GP is 0.98, which is inside the unit circle. Formally, a Dickey and Fuller (1979) stationarity test rejects the null of a unit root for $\log GP$ at the 5% level.[10] The high persistence of GP is in contrast to the view that shocks to gold prices are transient phenomena. Innovations in GP are uncorrelated with Pastor and Stambaugh (2003) liquidity factor innovations.

GP is countercyclical and strongly negatively correlated with equity valuation ratios; GP is high when stock prices are low. The strong positive correlation between GP and the default spread suggests that GP is high when firms with low credit ratings are more likely to default,

[10]The optimal number of lags is chosen based on the Ng and Perron (1995) sequential t-test.

which raises the required yield on their corporate bonds. GP is positively correlated with $ICC$ since the cost of capital for firms is high in adverse economic conditions. High values of $CAY$ are associated with high risk premia, and accordingly we see a positive correlation of GP with $CAY$ (Lettau and Ludvigson (2001)). GP is not correlated with $INFL$; this is expected, since inflation equally affects both the numerator and denominator of the ratio.

*1.2.2. Stock Return Predictability*

My measure of U.S. stock returns is the CRSP value-weighted index. The risk-free rate is the 1 month U.S. Treasury bill rate. I compare the performance of GP to various forecasting variables proposed in the literature. Table 1.2 shows the main predictability result of the paper. I run the regression:

$$\frac{12}{h}\sum_{i=1}^{h}\log R_{t+i} - \log R^f_{t+i} = \beta_0 + \beta_1 \log GP_t + \epsilon_{t+h} \tag{1.1}$$

Long-horizon returns are constructed from overlapping monthly returns. The top panel uses ordinary least squares regression with Newey and West (1987) HAC robust standard errors.[11] At the 1 month horizon, the degree of predictability is fairly low with an $R^2$ just above 1%; however, the estimated slope is statistically significant with a 2.82 t-statistic. We see similar patterns of predictability up to the 1 year horizon, which has an $R^2$ of 16.57%. The bottom panel uses the vector autoregression (VAR) framework as in Hodrick (1992), which is potentially more conservative for overlapping returns, although it imposes parametric assumptions. The point estimates are very similar, although the $R^2$ is lower (yet still very large at 10.89% for the 1 year horizon) using the VAR. For longer horizons of 2 to 5 years, the estimated coefficients are still significant although the magnitude is decreasing. The estimated coefficient on $\log GP$ for the one year horizon is 0.243, the standard deviation of $\log GP$ is 0.266, so a one standard deviation increase in $\log GP$ is associated with a 6.4% increase in U.S. stock market excess returns over the following year. For all horizons from

[11] I also conduct a robust check using Hodrick (1992) standard errors and confirm that the results are not sensitive to the choice of standard errors.

1 month to 5 years, the estimated slopes are statistically significant.

Table 1.3 shows the results of univariate predictability regressions for each of the predictors. For short horizon returns (1 and 3 months), only GP, and $VRP$ are statistically significant at conventional levels, with $ICC$ significant at the 10% level. At the intermediate 1 year horizon, GP, $ICC$, and $VRP$ are strongly significant, while $TMSP$ and $INFL$ are marginally significant. At this frequency, GP has the highest $R^2$ of all predictors. For long horizon (e.g. 5 year) returns, GP is still significant, while valuation ratios, $CAY$, and $TMSP$ are also significant.

How does GP stack up against other predictors in a horse race? The regression is:

$$\frac{12}{h}\sum_{i=1}^{h}\log R_{t+i} - \log R^f_{t+i} = \beta_0 + \beta_1 \log GP_t + \beta_2 X_t + \epsilon_{t+h} \tag{1.2}$$

where $X_t$ is another predictor. Table 1.4 shows the results for 1 and 3 month horizons. The first two columns under each return horizon refer to the coefficient and t-statistic, respectively, for $\beta_1$ in equation (1.2), and the next two columns are the coefficient and t-statistic for $\beta_2$.[12] At short horizons, shows $VRP$ is a strong predictor, with large t-statistics and high $R^2$. This result supports the findings of Bollerslev et al. (2009) and Drechsler and Yaron (2011) on more recent data. However, GP is still significant, even after controlling for $VRP$. The incremental $R^2$ at the 1 month horizon is 1.8%. Similar results hold for the 3 month horizon. Table 1.5 shows the results for long horizon returns. In most cases, GP drives out the significance of the other predictor, with the exception of $INFL$ at the 1 year horizon and $ICC$, $TMSP$, and $CAY$ (marginally) at the 5 year horizon. Including alternative predictors leaves the magnitude of the coefficient on GP largely unchanged. The evidence suggests that GP is robust to and in most cases outperforms other forecasting variables proposed in the literature.

Goyal and Welch (2008) argue that predictors such as the price-dividend ratio do not

[12]For bivariate regressions as well, the results using the Hodrick (1992) VAR methodology are similar.

perform well out-of-sample. I test out-of-sample robustness Out-of-Sample $R^2$. If GP is a robust predictor, Out-of-Sample $R^2$ should be significantly greater than zero and similar to in-sample counterparts. The statistic is given by:

$$R^2_{OS} = 1 - \frac{\sum_{k=1}^{T-m} (r^e_{m+k} - \widehat{r^e}_{m+k})^2}{\sum_{k=1}^{T-m} (r^e_{m+k} - \overline{r^e}_{m+k})^2} \tag{1.3}$$

We can calculate $R^2_{OS}$ using either an expanding window (use all data available from month 1 to month $m$, so the regression sample expands at each time step), or a rolling window of length $m$ (use only the past $m$ months of data at each time step). In both cases, I estimate equation (1.1) in the estimation period, compute the squared prediction error over the next period and increment my time step. An expanding window uses more available data, while a rolling window better accounts for potential time variation in the predictive relationship. I consider windows of length 120 months and 180 months to estimate betas, and predict the return in the next period. The p-values are from the Clark and West (2007) adjusted-MSPE statistic:

$$f_{t+1} = (r_{t+1} - \overline{r}_{t+1})^2 - \left[(r_{t+1} - \widehat{r}_{t+1})^2 - (\overline{r}_{t+1} - \widehat{r}_{t+1})^2\right]$$

which is regressed against a constant and the test is a one-sided test of whether $R^2_{OS} > 0$.

Table 1.6 shows the results from the out-of-sample analysis. With the exception of 1 month horizon rolling 10 year window regressions, all other combinations of forecast horizons and methods give large, positive, and significant out-of-sample $R^2$ values. The pattern of $R^2$ as we increase the forecast horizon are similar between rolling and expanding methods. Goyal and Welch (2008) find that for predictors such as the price-dividend ratio, the predictive ability is diminished in out-of-sample tests. For GP, the out-of-sample $R^2$ are significantly greater than zero and similar to the in-sample $R^2$. Figure 1.3 shows the estimated slopes and 95% error bands for both rolling and expanding methods with either 120 or 180 month

windows for 2 year-ahead predictive regressions. We see that the estimates are stable, never change signs, and are statistically significant in nearly all sub-samples. For comparison, Figure 1.4 plots the same sub-sample betas for $\log PD$. We see a lot of variation in the estimated coefficients with numerous sign changes and weak statistical significance in sub-samples (95% intervals straddle zero). The evidence suggests return predictability by GP is robust both out-of-sample and over sub-samples. In Appendix A.1.1, I show that GP is robust to finite-sample bias (Stambaugh (1999)) and size distortions (Torous, Valkanov, and Yan (2004)). In Appendix A.1.2, I show that realized utility gains are high for mean-variance investors using GP for portfolio allocation.

Gold and platinum are globally traded assets. This suggests that GP should also predict future stock returns in international markets. I run the same predictive regressions as in equation (1.1) using the MSCI World Index, which is a U.S. dollar denominated index composed of stocks from 23 Developed Markets countries covering approximately 85% of the free float-adjusted market capitalization in each country. Since the index is dollar denominated, I use the U.S. Treasury bill rate as the risk-free rate. Table 1.7 shows that the patterns of predictability are very similar to the U.S. results: high GP predicts high future excess returns, although the coefficients are somewhat smaller in magnitude than for U.S. returns. Since there may be some concern that the world portfolio consists of a large proportion of U.S. stocks, I also run the same predictability regressions for other developed countries. Panel B of Table 1.7 reports the results for the U.K., Switzerland, Japan, and Sweden. I use the MSCI country indices for each of these countries, denominated in the local currency. The risk-free rate is the local currency treasury bill rate. The results for the U.K., Switzerland, and Sweden are nearly the same as for the U.S., while Japan shows significant predictability in terms of the magnitude of estimated slopes, albeit smaller t-statistics (significant at the 10% level) and somewhat lower $R^2$. The results suggest that GP can predict equity risk premia for the U.S. market as well as international markets, which also mitigates potential concerns about data snooping (Ang and Bekaert (2007)).

### *1.2.3. Dividend Growth Predictability*

I have argued that stock return predictability by GP is driven by time variation in risk premia and not from news about future dividend growth rates. Some may argue that platinum has a characteristic not shared by gold: it is demanded by the automotive industry for catalytic converters. Is it possible that it is actually bad news about the future cash flows of car makers (GP is low when platinum is expensive, which is bad news for future cash flows of car makers) that drives the predictability through a cash-flow channel? I run standard dividend growth predictive regressions similar to Cochrane (2008) on real dividend growth rates ($\Delta d_t$) and real earnings growth rates ($\Delta e_t$):

$$\begin{aligned} \frac{12}{h}\sum_{i=1}^{h}\Delta d_{t+i} &= \beta_0 + \beta_1 \log GP_t + \epsilon_{t+h} \\ \frac{12}{h}\sum_{i=1}^{h}\Delta e_{t+i} &= \beta_0 + \beta_1 \log GP_t + \epsilon_{t+h} \end{aligned} \tag{1.4}$$

The results in Table 1.8 show no evidence of dividend growth predictability by GP.[13] For dividend growth, none of the estimated slopes from 1 year to 5 year horizons are statistically different from 0, and the $R^2$ are all nearly zero. For earnings growth, the $R^2$ are slightly higher but the t-statistics suggest the slopes are not significantly different from zero. This is evidence that the predictability I document arises because of variation in risk premia rather than dividend growth.

### *1.2.4. GP and the Cross-Section of Stock Returns*

I examine the implications of GP risk for the cross-section of stock returns. As seen earlier, GP is countercylical and increases in times of economic distress. Stocks with high, positive covariation with GP innovations are therefore a good hedge against adverse states of high economic risk and low asset valuations, which suggests that GP should command a negative market price of risk in the cross-section. I estimate the risk exposures (betas) for each asset

[13]The earnings data is from Robert Shiller's website.

$i = 1, ..., N$ from time-series regressions

$$R^e_{i,t+1} = c_i + \beta_{i,\Delta gp} \Delta \log GP_{t+1} + \epsilon_{i,t+1} \tag{1.5}$$

where $R^e_{i,t+1}$ is the excess return for portfolio $i$ and $\Delta \log GP_{t+1} = \log GP_{t+1} - \mathbb{E}_t [\log GP_t]$ is the innovation in GP.[14] The slope coefficient $\beta_{i,\Delta gp}$ represents the portfolio exposure of asset $i$ to GP risk. In order to estimate the cross-sectional market price of risk associated with GP, I run a cross-sectional regression of time-series average excess returns on the risk exposures

$$\mathbb{E}\left[R^e_{i,t+1}\right] = \text{cons} + \beta_{i,\Delta gp} \lambda_{\Delta gp} + \upsilon_i \tag{1.6}$$

which yields estimates of the market price of risk $\lambda_{\Delta gp}$. I use the standard cross-section of ten portfolios sorted on the book-to-market ratio and ten portfolios sorted on size as my test assets. The data is monthly from 1975 - 2013. Recall that GP is constructed without any information from equity markets, which rules out any mechanical relationship between GP risk and the cross-section of stock returns. Furthermore, the parsimonious one-factor model avoids many statistical issues present in asset pricing tests that can mechanically produce high explanatory power. Panel A of the Table 1.9 shows that the market price of GP risk is significantly negative. Panel B of the Table further shows that the portfolio returns are all significantly and negatively exposed to GP risk; equity returns decrease contemporaneously when GP increases. The one-factor model featuring only GP risk can explain over 60% of the cross-sectional variation in average returns. Figure 1.5 graphically depicts the strong negative relationship between average excess returns and risk exposures (Panel A), and good fit between realized and model-predicted excess returns (Panel B). The cross-sectional results suggest that investors are willing to pay a premium for assets which hedge against increases in GP; in other words, the high-risk states which investors dislike are those associated with high GP.

[14] The results using first differences are nearly identical.

*1.2.5. GP and Tail Risk*

The evidence so far suggests that 1) GP is countercyclical and increases in times of economic distress, 2) GP positively predicts future stock market excess returns, 3) GP risk is negatively priced in the cross-section, and 4) GP is high when the default spread is high, which is when firms with low credit ratings have higher probability of default. A plausible interpretation consistent with these results is that GP captures tail risk in the economy. This is broadly consistent with the findings of Manela and Moreira (2014), who use machine learning techniques to quantify tail risk (disaster concerns) from newspaper headlines: "gold" is one of the top words which explains variation in investors' tail risk concerns. GP is persistent, which is consistent with the evidence of persistent tail risk in Kelly and Jiang (2014). Options are an ideal way to measure tail risk because their convex payoff structure contains rich information about the tail distribution of returns. I extract tail risk measures from options markets and investigate the association between GP and tail risk.

Out-of-the-money (OTM) index put options protect against stock market crashes. The slope of the implied volatility curve, defined as the implied volatility of an OTM put minus the implied volatility of an at-the-money (ATM) put with the same maturity, is a measure of tail risk in the economy (Pan (2002)). In the data, the implied volatility curve slopes upward to the left since OTM puts are relatively more expensive (Rubinstein (1994)). I take the implied volatility curve from OptionMetrics and define $SLOPE_t^{\Delta}$ as the implied volatility for an OTM put option between $20\Delta$ to $40\Delta$, which I subtract from the implied volatility of an ATM put ($50\Delta$).[15] The encompassing regression is:

$$\underbrace{SLOPE_t^{\Delta}}_{\sigma_{t,IV}^{OTM,\Delta}-\sigma_{t,IV}^{ATM}} = \beta_0 + \beta_1 \log GP_t + \beta_2 \sigma_{t,IV}^{ATM} + \epsilon_t. \tag{1.7}$$

To control for potential dependence of the slope on the level of implied volatility, I also control for $\sigma_{t,IV}^{ATM}$ on the right hand side of (1.7). Panel A of Table 1.10 shows the results.

[15] $\Delta$ can be interpreted as the risk-neutral probability of expiring in-the-money. Lower $\Delta$ options are further out of the money.

We see that GP is significant for all definitions of the implied volatility slope, both by itself and after controlling for the level of ATM implied volatility. The magnitude of the coefficients as well as the t-statistics and $R^2$ increase as the OTM put is further out-of-the-money (more tail risk). An alternative measure of tail risk is the Bakshi et al. (2003) model-free risk-neutral skewness. Bakshi and Kapadia (2003) and Jurek (2014) use this measure as a proxy for crash risk; more negative skewness is associated with more crash risk. The results in Panel B of Table 1.10 are similar to the results in Panel A: high GP is associated with more negative risk-neutral skewness and GP is significant even after controlling for the risk-neutral variance.

## 1.3. Gold and Platinum Markets

I examine key aspects of gold and platinum markets, including sources of demand for each metal, the leasing markets, and return dynamics. Understanding the leasing markets is important because no-arbitrage implies that investors are indifferent between buying gold (platinum) or leasing gold (platinum) in perpetuity. Understanding the variation in rental income will be important for the economic model and to compute gold and platinum returns.

### *1.3.1. Sources of Demand*

Figure 1.6 shows the annual percentage demand for gold (top panel) and platinum (bottom panel) for each of its major uses.[16] From 1990 - 2013, approximately 70% of gold demand was for jewellery, 15% for uses in technology (semiconductors, electronics), and 15% for investments (coins, bars, ETF inventory building). Over the same period, approximately 40% of platinum demand was for jewellery, 15% for technology, while only a small fraction (less than 5%) was demanded for investment purposes. Quite conspicuously, the biggest difference between the two metals comes from the 40% of platinum demand used by the automotive industry as catalytic converters to reduce emissions in automobiles (autocatalysts).[17] Demand for platinum as autocatalysts was spurred by clean air legislation in

[16]The data for gold is from Thomson-Reuters GFMS, and the data for platinum is from Johnson Matthey.

[17]While beyond the scope of this paper, Black (2000) (Chapter 5) provides an overview of the process by which platinum group metals catalyse the oxidation of hydrocarbons and carbon monoxide from internal

the 1970s - securing sufficient supplies of platinum at stable prices became essential for car makers. Black (2000) (Chapter 6) describes the long-term arrangements made between platinum producers and car makers:

> *"With the introduction of autocatalysts...producers entered into long term supply contracts with the auto manufacturers. Prices were negotiated on contracts lasting up to five years".*

The private sale of platinum directly from producers to car makers means the amount of platinum used in auto production does not enter the market. Therefore, net autocatalyst demand (in excess of salvage) acts as negative platinum supply shocks since it reduces the effective supply of platinum available for other uses. Under this view, the major source of demand for gold and platinum comes from the jewellery industry.[18]

### *1.3.2. Lease Rates*

Not surprisingly, jewellers are among the most active borrowers of gold and platinum. The LBMA and LPPM describe the leasing market:

> *"The inventory loan is the basic financial tool of the precious metals fabricating [industry]. For example, jewellery manufacturers can finance the raw material in their production process by leasing gold...The same kind of strategy would, for example, be adopted in platinum".*

Leasing is a convenient form of inventory financing widely practiced in both gold and platinum fabrication industries (LBMA and LPPM (2008)). Le and Zhu (2013) find that over the 1991 - 2007 sample, which purposely excludes the 2008 financial crisis to focus on normal time dynamics, gold lease rates are increasing in stock market returns. This is consistent with Aït-Sahalia, Parker, and Yogo (2004) who document strong positive covariation between stock returns and demand for luxury goods. In normal times, gold lease rates are

combustion engines.

[18] I use the term "jeweller" to refer to gold and platinum fabricators.

procyclical: as stock returns go up, jewellers have increased need for raw materials to meet high demand for finished products and increase their gold borrowing, which drives up gold lease rates.

The picture is different in times of economic distress. Figure 1.7 plots annualized gold lease rates from 2007 - 2009. While gold lease rates are about 1% on average, the cost of borrowing gold during the financial crisis jumped up threefold and high gold lease rates persisted throughout the crisis. This is much greater than the observed decline in gold prices during this period, which implies the rental income (economic value of holding gold) must have been very high during the crisis. Several factors lead to countercyclical behavior of lease rates in bad times. In severe economic conditions, lenders fear default by borrowers and decrease the supply of loans, which increases the cost of leasing precious metals (LBMA (2009)). Furthermore, to the extent that there are greater countercyclical benefits to service flows from gold relative to platinum in bad times (for reasons discussed in the introduction), expected gold rental income will exceed platinum. Risk premia are high in bad times, which raises discount rates and lowers the prices of stocks, gold, and platinum. However, since gold and platinum prices are equal to the discounted sum of future rental income, the increase in expected gold rental income cushions the fall in gold prices relative to platinum prices.

### *1.3.3. Gold and Platinum Returns*

Previous studies of gold returns (see e.g., Erb and Harvey (2013), Barro and Misra (2013)) focus only on the price appreciation of gold and do not include the rental income over the ownership period. While Barro and Misra (2013) are correct in stating that gold dividends are not directly observable given spot prices alone, we can compute and monetize rental income through the futures market.

Gold and platinum futures data comes from the commodities division (COMEX) of the CME (formerly NYMEX).[19] As is standard in the literature, I ignore mark-to-market of futures,

[19]The data is obtained from the Commodities Research Bureau (CRB) and are daily settlement prices direct from the exchange.

and also the delivery options embedded in futures with physical settlement. I assume that futures contracts will roll in the first week of the expiration month; it is estimated that only 1% to 2% of commodities futures contracts are actually delivered, so this approach should not result in too much measurement error (Hirschey and Nofsinger (2008), Chapter 19). I examine the resulting contract maturities and verify similar to Schwartz (1997) that the maturities are relatively constant. The lease rate is given by:

$$\text{Lease rate} = \text{Libor rate - Futures premium.} \tag{1.8}$$

Table 1.11 provides a table of cash flows analysis of the above from the perspective of a jeweller, who is a typical borrower in the leasing market. For my analysis, I use futures contracts closest to 3 months to maturity, and match it with the 3 month Libor rate to calculate the lease rate, which I then annualize.[20] I choose 3 month maturities to get a contract with high liquidity, short time-to-maturity, yet not too short so that the physical delivery option does not affect prices too much. I use average daily futures prices as the monthly futures price, since my measure of monthly spot prices are average daily spot prices over the month. Real gold and platinum returns (inclusive of rental income) are calculated in the standard way:

$$\begin{aligned} R_{t+1}^{real,\ gold} &= \frac{\left(\frac{P_{t+1}^{g}}{CPI_{t+1}} + \frac{D_{t+1}^{g}}{CPI_{t+1}}\right)}{\left(\frac{P_{t}^{g}}{CPI_{t}}\right)} \\ R_{t+1}^{real,\ platinum} &= \frac{\left(\frac{P_{t+1}^{x}}{CPI_{t+1}} + \frac{D_{t+1}^{x}}{CPI_{t+1}}\right)}{\left(\frac{P_{t}^{x}}{CPI_{t}}\right)} \end{aligned} \tag{1.9}$$

where $P_t^g$ is the gold price, $P_t^x$ is the platinum price, $D_t^g$ is the gold rental income at time $t$, and $D_t^x$ is the platinum rental income at time $t$, and $CPI_t$ is the consumer price index.[21] The returns to gold and platinum can be interpreted from the perspective of an investor who owns gold or platinum, and continuously leases the metal out, earning the rental income

[20] Prior to 1986, I use Eurodollar deposit rates.

[21] I use superscript $g$ to refer to gold, and superscript $x$ to refer to platinum.

and any price appreciation. The results are summarized in Table 1.12. Average gold excess returns are 2.40% per year, real gold return volatility is 16.76%, implying a Sharpe ratio of 0.14. Average platinum excess returns are 6.51% per year, real platinum return volatility is 22.18%, implying a Sharpe ratio of 0.29. For comparison, over the same period, the average excess return for U.S. equities is 7.53% per year, real equity return volatility is 15.11%. The gold risk premium is substantially lower than the equity risk premium. The risk premium for platinum is slightly lower than equities as well, although the volatility is higher. Gold lease rates are 1% per year on average. For comparison, Casassus and Collin-Dufresne (2005) estimate the gold lease rate to be 0.9% per year, while Le and Zhu (2013) find an average lease rate of about 1%. My estimate of the average platinum lease rate is 3.47% per year. The economic model must match the low risk premium, high volatility, and low lease rate of gold. At the same time, the model must also capture the relatively high risk premium, high volatility, and high lease rate of platinum, while fitting the asset pricing dynamics of equity markets and quantitatively accounting for the time variation and stock return predictability of GP observed in the data.

## 1.4. Economic Model

### *1.4.1. Economic Environment*

I analyze whether a general equilibrium model featuring time-varying disaster risk (Wachter (2013)) and shocks to preferences for gold and platinum can jointly explain the empirical facts documented in the previous sections. I assume an endowment economy with complete markets and an infinitely-lived representative investor with Duffie and Epstein (1992) stochastic differential utility, which is the continuous-time analog of Kreps and Porteus (1978) and Epstein and Zin (1989) recursive preferences. Recursive preferences allow for a separation between risk aversion and the intertemporal elasticity of substitution (IES).[22]

[22] A number of studies such as Bansal and Yaron (2004) argue that the IES should be greater than one. Others, such as Hall (1988) estimate IES to be significantly less than one, although time-varying consumption volatility can lead to large downward biases in the estimates of IES using the methodology employed in Hall (1988).

I focus on the case of unit IES, which is done both for tractability, and consistent with evidence in Vissing-Jørgensen (2002) and Hansen, Heaton, Lee, and Roussanov (2007).

Aggregate consumption growth is given by

$$d\log C_t = \bar{g}_c dt + \sigma_c dW_t^c + J_t^c dN_t^c \tag{1.10}$$

where $W_t^c$ is a standard Brownian motion and $N_t^c$ is a Poisson process whose intensity $\lambda_t$ is given by a Cox, Ingersoll, and Ross (1985) square root process

$$d\lambda_t = \kappa_\lambda(\xi_t - \lambda_t)dt + \sigma_\lambda\sqrt{\lambda_t}dW_t^\lambda + J_t^\lambda dN_t^\lambda \tag{1.11}$$

where $W_t^\lambda$ a standard Brownian motion and $N_t^\lambda$ is a Poisson process whose intensity is given by $\lambda_t^\lambda = \lambda_t$.[23] Drechsler and Yaron (2011) use a similar framework to model jumps in expected consumption growth and volatility. Allowing $\lambda_t$ to jump allows stock prices and volatility in the model to jump as well.[24] This also allows the model to explain the large observed jumps in volatility during the 2008 financial crisis as well as jumps in GP seen in Figure 1.2.[25] I solve for the stationary mean of $\lambda_t$ in Appendix A.1.3. $\lambda_t$ can be approximately thought of as the probability of a consumption disaster.[26] In the model, market volatility is endogenously determined, and evidence from the volatility estimation literature argues in favor of multiple time scales in volatility allowing for both long and short run components.[27] Also, as Seo and Wachter (2014) demonstrate, a one-factor model without time-variation in the long-run mean of $\lambda_t$ generates the counterfactual prediction that the slope of the implied volatility curve decreases as the disaster intensity increases. This arises because stock return volatility is endogenously determined and is driven by $\lambda_t$

[23] Nowotny (2011) considers the implications of self-exciting intensity processes to model persistent disaster states. My setup differs since realized jumps in consumption do not trigger increases in $\lambda_t$.

[24] This is consistent with the evidence in Duffie et al. (2000), Broadie et al. (2007), Eraker and Shaliastovich (2008), and Tauchen and Todorov (2011).

[25] For parsimony, I do not distinguish explicitly between $X_{t-}$ and $X_t$ in my notation, as it should be clear from the context.

[26] The probability of $k$ jumps over an interval of time $\Delta t \approx e^{\lambda_t \Delta t}\frac{(\lambda_t \Delta t)^k}{k!}$.

[27] See e.g., Alizadeh, Brandt, and Diebold (2002) and Chernov, Gallant, Ghysels, and Tauchen (2003).

itself. To relieve this tension, and consistent with the evidence from the volatility estimation literature, I follow Seo and Wachter and allow the long-run mean of $\lambda_t$ to be a stochastic process $\xi_t$, which itself follows a square root process

$$d\xi_t = \kappa_\xi(\bar{\xi} - \xi_t)dt + \sigma_\xi\sqrt{\xi_t}dW_t^\xi \tag{1.12}$$

where $W_t^\xi$ is a standard Brownian motion. All Brownian motions and Poisson processes are assumed to be independent.

The size of the consumption jump, $J_t^c$ is drawn from the multinomial disaster distribution of Barro and Ursua (2008), using data obtained from Robert Barro's website. While the modeling paradigm uses the rare disasters framework, the disasters I have in mind are smaller. This will be made clearer in the following section when I discuss the model calibration. The size of the jump in $\lambda_t$ is given by $J_t^\lambda$, which follows an exponential distribution with mean $\mu_\lambda$. Equity is modeled as a leveraged claim on aggregate consumption following Abel (1999). The aggregate dividend at time $t$ is $D_t = C_t^\phi$, for leverage parameter $\phi$, which implies that dividend growth dynamics are given by

$$d\log D_t = \phi\bar{g_c}dt + \phi\sigma_c dW_t^c + \phi J_t^c dN_t^c. \tag{1.13}$$

### *1.4.2. Gold and Platinum Supply*

Gold and platinum do not depreciate, and consumption of the service flow from the stock of gold and platinum today does not render it less capable of providing the same service flow tomorrow. I model gold and platinum as non-depreciating durable goods. This means that the time $t$ aggregate stock of gold and platinum increase one-to-one with the time $t$ increment (accumulation) to the stock (Cuoco and Liu (2000)). In Appendix A.1.4, I use data from world gold and platinum mine production to establish the properties of the gold and platinum endowment processes. The key stylized facts are: 1) the log growth rates of

the aggregate per-capita gold and platinum stocks are smooth with no evidence of disasters, and 2) the aggregate per-capita gold and platinum stocks are cointegrated.[28] Given these facts, I model $\log G_t$ (the aggregate stock of gold) using a simple geometric Brownian motion which is not subject to disasters. Consistent with the empirical evidence, $\log G_t$ and $\log X_t$ (the aggregate stock of platinum) are modeled as cointegrated processes so that $\log X_t - \log G_t = \log Z_t$ is a stationary process which itself follows an Ornstein-Uhlenbeck process with long-run mean $\mu_z$ and reversion parameter $\theta_z$:

$$\begin{aligned} d\log G_t &= \mu_g dt + \sigma_g dW_t^g \\ d\log Z_t &= \theta_z\left(\mu_z - \log Z_t\right)dt + \sigma_z dW_t^z \\ \log X_t &= \log G_t + \log Z_t. \end{aligned} \tag{1.14}$$

All parameters for the gold and platinum supply dynamics are directly estimated from the data.

*1.4.3. Preferences*

The representative investor's utility function is defined recursively as

$$\begin{aligned} V_t &= \mathbb{E}_t\left[\int_t^\infty f(\Omega_s, V_s)ds\right] \\ \text{for } f(\Omega, V) &= \delta(1-\gamma)V\left[\log\Omega - \frac{1}{1-\gamma}\log\left(1-\gamma\right)V\right] \\ \text{and } \Omega_t &= \left[C_t^{1-\frac{1}{\epsilon}} + \alpha_t G_t^{1-\frac{1}{\epsilon}} + \beta_t X_t^{1-\frac{1}{\epsilon}}\right]^{\frac{1}{1-\frac{1}{\epsilon}}} \end{aligned} \tag{1.15}$$

where $f(\Omega, V)$ describes the trade-off between current consumption $\Omega_t$ and the continuation utility $V_t$. The subjective time preference parameter is $\delta$, and $\gamma$ is commonly interpreted as the coefficient of relative risk aversion.

The consumption aggregator $\Omega_t$ is a constant elasticity of substitution (CES) aggregator

[28] Barro and Misra (2013) also find no evidence of disasters in the per-capita gold stock, using data since 1836.

over nondurable consumption $C_t$, the gold stock $G_t$, and the platinum stock $X_t$.[29] The intratemporal elasticity of substitution is $\epsilon$.[30]

The processes $\alpha_t$ and $\beta_t$ capture in reduced-form time-varying preferences for gold and platinum:

$$\begin{aligned} \alpha_t &= \exp(a_1 + a_2\lambda_t) \\ \beta_t &= \exp(b_1 + b_2\lambda_t). \end{aligned} \tag{1.16}$$

Specifically, $\alpha_t$ and $\beta_t$ represent the relative importance of gold and platinum service flows in the intratemporal consumption aggregator. Preference for precious metals responds to changes in $\lambda_t$ but not directly to $\xi_t$, since $\lambda_t$ is the probability of a consumption disaster. While the processes $\alpha_t$ and $\beta_t$ gives me some additional flexibility, they depend completely on existing state variables and no new state variables are being added. The parameter $a_2$ (and $b_2$) cannot be arbitrarily set. We want a relatively high value of $a_2$ to generate enough countercyclical dynamics to match the low observed gold risk premium. However, when $a_2$ is too big, gold return volatility becomes too low, and gold lease rates will also be too low. Additionally, existence of solutions for gold and platinum price-dividend ratios places restrictions on the maximum $a_2$ and $b_2$ allowed, and this bound jointly depends on model parameters such as the volatility and persistence of state variables, the severity of jumps, and risk aversion. Changing these parameters to allow for high $a_2$ will affect equity market dynamics as well.[31]

[29]The agent derives utility from gold and platinum service flows in direct proportion to its stock. This is a standard way to model preference for multiple types of goods, which has been used in the durable goods literature by Ogaki and Reinhart (1998) and Yogo (2006).

[30]I use a CES aggregator with the same elasticity of substitution across all pairs of goods for parsimony and tractability relative to a specification with nested CES aggregators and separate elasticities.

[31]I have solved a version of the model based on the long-run risks (LRR) framework pioneered by Bansal and Yaron (2004) featuring jumps to uncertainty as in Eraker and Shaliastovich (2008) and Drechsler and Yaron (2011). The model is also able to deliver many similar results if the countercyclical components $\alpha_t$ and $\beta_t$ are allowed to load on all state variables. The LRR framework delivers dividend growth predictability by GP, which is not seen in the data. I have opted for a stochastic disaster risk framework mostly for parsimony and to avoid these issues.

*1.4.4. Asset Pricing*

Duffie and Skiadas (1994) show that

$$\pi_t = \exp\left(\int_0^\infty f_V(\Omega_s, V_s)d\right)\ f_\Omega(\Omega_t, V_t)$$

can serve as the state-price density in this economy. In equilibrium, the relationship between $V_t$ and the state variables is given by

$$V_t = \frac{C_t^{1-\gamma}}{1-\gamma} e^{a+b_\lambda \lambda_t + b_\xi \xi_t}.$$

In this economy, the equation for the state price density becomes

$$\pi_t = \exp(\eta t - \delta b_\lambda \int_0^t \lambda_s ds - \delta b_\xi \int_0^t \xi_s ds)\ \delta\ \Omega_t^{-\gamma} e^{a+b_\lambda \lambda_t + b_\xi \xi_t}$$

where $\eta = -\delta(a+1)$ and $a,\ b_\lambda,\ b_\xi$ are the solutions to a system of equations given in Appendix A.1.5. Following Barro and Misra (2013), I assume that outlays on gold and platinum are negligible relative to nondurable consumption, which implies that $\Omega_t \approx C_t$. Under this assumption, the state price density is given by

$$\pi_t \approx \exp(\eta t - \delta b_\lambda \int_0^t \lambda_s ds - \delta b_\xi \int_0^t \xi_s ds)\ \delta\ C_t^{-\gamma} e^{a+b_\lambda \lambda_t + b_\xi \xi_t} \tag{1.17}$$

The levels of $\alpha_t$ and $\beta_t$ are small because per-capita expenditures on gold and platinum are small compared to expenditures on nondurable goods and services. When the CES aggregator is over multiple sources of consumption with large expenditure shares, such as nondurable and durable consumption or housing, this approximation will become wildly inaccurate; for example, Gomes, Kogan, and Yogo (2009) estimate the expenditure share of durable goods to be 50%, in which case this assumption would not be innocuous. In economic terms, the assumption implies that shocks to the supply of gold and platinum are unpriced. A mine shutdown in South Africa, for example, would affect gold and platinum

prices, but would conceivably not affect aggregate stock market risk premia, which seems economically plausible. Going forward, I will assume that the approximation is accurate and describe dynamics of the stochastic discount factor in (1.17) with an equality sign.

The instantaneous risk-free rate is given by

$$r_t^f = \delta + (\bar{g_c} + \frac{1}{2}\sigma_c^2) - \gamma\sigma_c^2 + \lambda_t \mathbb{E}_v \left[ e^{(1-\gamma)J_t^c} - e^{-\gamma J_t^c} \right].$$

I follow Barro (2006) and Wachter (2013) and suppose that if a disaster occurs, the government will default on debt obligations with probability $q$, leading to a loss in the same proportion as the consumption loss in the disaster.

The user costs (rental income) of gold and platinum are determined in equilibrium by the intratemporal optimality conditions:

$$\begin{aligned} Q_{g,t} = \frac{\Omega_G}{\Omega_C} = & \underbrace{\alpha_t \times \left(\frac{C_t}{G_t}\right)^{\frac{1}{\epsilon}}}_{\text{countercyclical} \times \text{procyclical}} \\ Q_{x,t} = \frac{\Omega_X}{\Omega_C} = & \underbrace{\beta_t \times \left(\frac{C_t}{X_t}\right)^{\frac{1}{\epsilon}}}_{\text{countercyclical} \times \text{procyclical}} \end{aligned} \tag{1.18}$$

where $Q_{g,t}$ is the user cost of gold and $Q_{x,t}$ is the user cost of platinum. Notice from equation (1.18) that the intratemporal elasticity of substitution $\epsilon$ behaves like the inverse of the leverage parameter $\phi$, since shocks to gold and platinum supply are small and unpriced. When $\frac{1}{\epsilon} < \phi$, gold and platinum will be safer than levered equity and command a lower risk premium, while the opposite will be true if $\frac{1}{\epsilon} > \phi$. Lower values of $\epsilon$ lead to higher risk premia and volatility for gold and platinum returns, and also imply greater complementarity between nondurable consumption, gold, and platinum. Barro and Misra (2013) set $\epsilon > 1$, which makes gold less risky than unlevered equity. The authors use this mechanism to generate a low gold risk premium. However, as Wachter (2013) points out, under recursive preferences, when $\phi < 1$ (in this case, $\epsilon > 1$) the price-dividend ratio is increasing $\lambda_t$.

The same result holds in my model since gold and platinum supply shocks are unpriced. This means that under the Barro and Misra (2013) assumption that $\epsilon > 1$, the model would predict that gold lease rates fall (gold prices rise) when the probability of a disaster increases, which is counterfactual in light of Figures 1.1 and 1.7. Intuition suggests $\epsilon < 1$ is more reasonable if we view gold and platinum as jewellery, since jewellery complements nondurable consumption but does not substitute for it. Furthermore, $\epsilon > 1$ results in gold return volatility being too low because in this case gold becomes a deleveraged consumption claim. In my calibration, I set $\epsilon = \frac{1}{\phi}$ so that all the countercyclical properties of gold and platinum arise through $\alpha_t$ and $\beta_t$.

Let $P_t$ be the price of a claim to the stream of dividends $D_t$, and $P_t^{t+\tau}$ be the price of the asset which pays the single risky dividend $D_{t+\tau}$ and nothing else. No arbitrage implies that $\pi_t P_t^{t+\tau}$ is a martingale, which implies that the equity price-dividend ratio is given by

$$\frac{P_t}{D_t} = \int_0^\infty e^{a_\phi(\tau) + b_\phi(\tau)\lambda_t + c_\phi(\tau)\xi_t} d\tau = G(\lambda_t, \xi_t). \tag{1.19}$$

Similar arguments hold for $P_{g,t}$ and $P_{x,t}$, which are the claims to gold and platinum, respectively:

$$\begin{aligned} \frac{P_{g,t}}{Q_{g,t}} &= \int_0^\infty e^{a_g(\tau) + b_g(\tau)\lambda_t + c_g(\tau)\xi_t} d\tau = G^g(\lambda_t, \xi_t) \\ \frac{P_{x,t}}{Q_{x,t}} &= \int_0^\infty e^{a_x(\tau) + b_x(\tau)\lambda_t + c_x(\tau)\xi_t + d_x(\tau)\log Z_t} d\tau = G^x(\lambda_t, \xi_t, \log Z_t). \end{aligned} \tag{1.20}$$

The equity functions $a_\phi(\tau)$, $b_\phi(\tau)$, $c_\phi(\tau)$, gold functions $a_g(\tau)$, $b_g(\tau)$, $c_g(\tau)$, and platinum functions $a_x(\tau)$, $b_x(\tau)$, $c_x(\tau)$, $d_x(\tau)$ are given by the solution to systems of ordinary differential equations described in Appendix A.1.5.

*1.4.5. GP in the Model*

While I use the exact log GP ratio in my model simulations, a log-linearization conveys the economic intuition more clearly.[32]

In Appendix A.1.6, I show that we can write log-linearized gold ($P_{g,t}$) and platinum ($P_{x,t}$) prices as

$$\begin{aligned}\log P_{g,t} &= A_g + \frac{1}{\epsilon}\log C_t - \frac{1}{\epsilon}\log G_t + \underbrace{(a_2 + b^*_{g,\lambda})}_{<0}\lambda_t + \underbrace{b^*_{g,\xi}}_{<0}\xi_t \\ \log P_{x,t} &= A_x + \frac{1}{\epsilon}\log C_t - \frac{1}{\epsilon}\log G_t + \underbrace{(b_2 + b^*_{x,\lambda})}_{<0}\lambda_t + \underbrace{b^*_{x,\xi}}_{<0}\xi_t + \underbrace{(b^*_{x,Z} - \frac{1}{\epsilon})}_{<0}\log Z_t\end{aligned} \tag{1.21}$$

where $A_g, A_x, b^*_{g,\lambda}, b^*_{g,\xi}, b^*_{x,\lambda}, b^*_{x,\xi}, b^*_{x,Z}$ are constants described in Appendix.1.6. Positive shocks to $\log C_t$ imply higher service flows and raise gold and platinum prices. The increase is greater than the increase in consumption itself because of complementarity between non-durable consumption and gold and platinum service flows ($\frac{1}{\epsilon} > 1$). High $\log G_t$ lowers gold prices since the quantity of gold becomes less scarce, and also lowers platinum prices due to cointegration. Higher $\log Z_t$ means that (all else equal) the quantity of platinum is less scarce, which also lowers platinum prices. Under my model calibration, strong discount rate effects imply that, despite $a_2, b_2 > 0$, the overall response of gold and platinum to increases in $\lambda_t$ and $\xi_t$ are negative, so that gold and platinum prices fall as disaster risks increase.

The log GP ratio is the difference between the log gold and platinum prices and is given by

[32]The exact log GP ratio is given by

$$\begin{aligned}\log GP_t = \log\frac{P_{g,t}}{P_{x,t}} &= \log\frac{G^g(\lambda_t,\xi_t)}{G^x(\lambda_t,\xi_t,\log Z_t)}\frac{Q_{g,t}}{Q_{x,t}} \\ &= (a_1 - b_1) + \log\frac{G^g(\lambda_t,\xi_t)}{G^x(\lambda_t,\xi_t,\log Z_t)} + (a_2 - b_2)\lambda_t + \frac{1}{\epsilon}\log Z_t.\end{aligned}$$

$$\begin{aligned}\log GP_t &= \log \frac{P_{g,t}}{P_{x,t}} \\ &= \text{cons} + \underbrace{(\frac{1}{\epsilon} - b^*_{x,Z})}_{>0} \log Z_t + \underbrace{(a_2 - b_2 + b^*_{g,\lambda} - b^*_{x,\lambda})}_{>0} \lambda_t + \underbrace{(b^*_{g,\xi} - b^*_{x,\xi})}_{>0} \xi_t.\end{aligned} \tag{1.22}$$

Shocks to $\log C_t$ (which can be thought of as shocks to jewellery demand) affect gold and platinum prices equally, leaving GP insulated from consumption shocks. Likewise, shocks to $\log G_t$ alone also cancel out and only the relative difference in supply $\log Z_t$ matters for GP. Platinum is more expensive than gold on average because $\log X_t < \log G_t$ on average (platinum is more scarce). When $\log Z_t$ goes up, gold becomes scarce relative to platinum, which increases GP. In the model, GP is increasing in both $\lambda_t$ and $\xi_t$.[33] High disaster probabilities imply high risk premia, which leads to high discount rates and low equity prices. Since $a_2$ and $b_2$ are positive, the service flows from gold and platinum increase when disaster probabilities increase, which partially offsets the higher discount rates and cushions the fall in prices. This works similar to a cash flow effect, where the cash flow represents gold and platinum rental income. Furthermore, $a_2 > b_2$ implies that the higher service flow is greater for gold relative to platinum, which not only affects the immediate service flow but also expected future service flow (rental income) through persistence in disaster probabilities. This means that gold and platinum prices both fall as disaster probabilities increase, but gold prices fall by less relative to platinum and GP is increasing in the disaster probabilities. The fact that GP increases in $\lambda_t$ and $\xi_t$ allows the model to generate the observed return predictability at both long and short horizons.

The $\log Z_t$ term is not priced by the stochastic discount factor but does affect the volatility and persistence of GP. Stationarity of GP in the model is assured because $\log Z_t$ is stationary, or in other words, because $\log G_t$ and $\log X_t$ are cointegrated. Interestingly, while shocks to $\log Z_t$ affect GP, they do not affect return predictability, which suggests that controlling

[33] That the GP ratio increases in both $\lambda_t$ and $\xi_t$ is dependent on the calibration. This holds under the model parameters I use for this model.

for $\log Z_t$ in the data can potentially lead to even stronger return predictability by GP. I verify that this indeed holds in the data and discuss the results in the following section.

## 1.5. Calibration and Model Simulation Results

My parameter choices are given in Table 1.13. I have opted for smaller average jump sizes with an average disaster size of 15%. Barro (2006) uses the dataset of Madison (2003) and found the average disaster size to be 29%. Barro and Ursua (2008) update Madison (2003) and find that the average disaster size is between 21-22%; this disaster distribution is also used in Wachter (2013).[34] I opt for smaller average disaster sizes in line with evidence from Nakamura, Steinsson, Barro, and Ursua (2013), who document partial recoveries after disasters, and estimate the average permanent impact of disasters to be about 15%. While the actual probability of these smaller disasters is 5.85%, I opt for a more conservative calibration of 4%, which is achieved using a $\bar{\xi} = 0.0355$ as in Wachter (2013) along with an average jump size of $\mu_\lambda = 0.03$ in the event of a jump in $\lambda_t$. Figure 1.8 compares my multinomial jump size distribution with smaller average jump sizes to the distribution used in Barro and Ursua (2008) and Wachter (2013). An important challenge in calibrating representative investor models is to match the high observed volatility of the price-dividend ratio. The model places an upper bound on the amount of volatility in the state variables that can be allowed for solutions to exist (this is clearly seen in the equations for the Epstein-Zin discount factor in Appendix A.1.5). I fix $\sigma_\xi$ such that the discriminant in the solution to $b_\xi$ is zero, which helps match the high volatility of the price-dividend ratio and also reduces the number of free parameters. The $\lambda_t$ process is calibrated to be less persistent than $\xi_t$.

Table 1.14 describes the fit of the model to the data. State variables are simulated at a monthly frequency and aggregated to an annual frequency. The data moments are from 1975-2013. The model matches the low gold risk premium, relatively high gold return

[34]The cutoff in Barro (2006) for a disaster was a 15% peak-to-trough decline in GDP per capita, while Barro and Ursua (2008) used a cut-off of 10%. To achieve an average disaster size of 15%, my cutoff is 6%.

volatility, low Sharpe ratio, and low lease rate. The model-implied gold lease rate is 0.93%, which compares well to the 1% lease rate in the data. Lease rates in the model are the convenience yield, which corresponds to the dividend yield (dividend over price). For comparison, I also present the model 90% confidence intervals for simulation paths in which no disaster occurred. While these no-disaster intervals are more appropriate to compare against stock and bond moments (since no disasters have occurred in the recent U.S. data, on which the stock and bond returns are based), for gold and platinum returns it is more natural to compare against population moments, since there have been 21 economic disasters from 1975 - 2006 in international markets (using my disaster cutoff) based on the Barro and Ursua (2008) dataset (including several OECD countries), which can conceivably affect gold and platinum returns and volatilities. The model explains the expected returns, volatilities and lease rates for platinum as well, including the high lease rate and high volatility. The model also accounts for time variation in GP, with the volatility and persistence of GP falling right inside the 90% confidence intervals. The median persistence for all simulations matches the data estimate nearly perfectly. Following this, I run the below return predictability regressions using model excess stock returns and GP:

$$\frac{1}{h}\sum_{i=1}^{h}\log(R^e_{t+i}) - \log(R^b_{t+i}) = \beta_0 + \beta_1 \log(GP_t) + \epsilon_{t+h}.$$

The left hand side is the normalized excess return for one year up through five years ahead, while the right hand side is the model GP. The results are shown in the top panel of Table 1.15. The data estimates fall right in the model confidence intervals, with the data $R^2$ estimates very close to the median values.[35] Thus, the model can explain the observed predictability of returns by GP. Similar to the data, the model delivers very low to negligible dividend growth predictability, similar to (Wachter (2013)). The model can also account for the observed relationship between GP and the slope of the implied volatility curve for index options, as detailed in Appendix A.1.7.

[35]It is difficult to decide which, all simulations or no disasters, is most appropriate for the predictability exercise, since U.S. stock returns were not affected by domestic disasters, while the GP ratio is potentially affected by international disasters. For completeness, I include both sets of results.

How well have I captured the effect of supply dynamics on GP? Is there predictability coming from the supply effects (including autocatalyst demand)? The second panel of Table 1.15 investigates this issue. I regress GP on $\log Z_t$ inside the model, and we see that the data estimate falls right inside the 90% interval.[36] Since the leading coefficient on $\log Z_t$ in the model depends on $\frac{1}{\epsilon}$, this serves as a further check on the assumed complementarity ($\epsilon < 1$) between jewellery (gold and platinum) and nondurable consumption. Under a calibration where $\epsilon > 1$ as in Barro and Misra (2013), this regression in the data results in a coefficient smaller than 1. The second regression in this panel investigates return predictability by $\log Z_t$ in both the model and the data. In the model, $\log Z_t$ does not predict returns by construction, although in small samples it is occasionally possible to spuriously find weak evidence of predictability. Both the population and median values, however, show that there is no predictability coming from the supply channels. I run the same regression in the data and find no evidence of predictability through $\log Z_t$, which is further evidence that the predictability does not come from a cash flow channel.

These results for repeated samples of 39 years lead to an interesting finding. Time-variation in GP over finite samples is affected by $\log Z_t$, which is not a priced variable in this economy. The third panel shows return predictability regressions where I control for the effect of $\log Z_t$, which adds volatility and persistence to GP without adding predictive power. We see in this case that the point estimates increase at all horizons, and now the 90% interval for return predictability by GP does not contain 0. The $R^2$ increase over all horizons quite dramatically. In the data, we can separately identify $\log Z_t$ and $\log A_t$, the aggregate per-capita stock of platinum used as autocatalysts. Empirically, a regression of GP on $\log Z_t$ gives a significant, positive coefficient while a regression of GP on $\log A_t$ gives a significant, negative coefficient. When $\log Z_t$ is high, platinum is relatively more plentiful (gold more scarce) so gold is relatively more expensive than platinum. High values of $\log A_t$ correspond to high demand for platinum as autocatalysts, which is associated with higher platinum prices (lower GP). Neither $\log Z_t$ nor $\log A_t$ seem to predict returns in the data. Controlling

[36] For the data $\log Z_t$, I interpolate annual values to monthly values.

for persistent supply effects, the persistence of GP is lower; the monthy AR(1) coefficient is 0.962, which implies a half-life of about 1.5 years.

## 1.6. Conclusion

The risk and return tradeoff is one of the central tenets of asset pricing theory, yet empirically identifying a viable proxy for risk, manifest through robust return predictability, cross-sectional pricing, and basic economic intuition, has been largely elusive in the literature. In this paper, I show that the ratio of gold to platinum prices (GP) proxies for an important aggregate source of risk in the economy. GP predicts future stock returns in the time-series, outperforms other predictors proposed in the literature, and GP risk is priced in the cross-section of stock returns. GP is persistent and significantly correlated with tail risk measures implied by options markets. An equilibrium model with time-varying tail risk and shocks to preferences for gold and platinum can quantitatively account for the asset pricing dynamics of equity, gold, and platinum markets, as well as the time variation and return predictability of GP. In the model, higher aggregate risk lowers gold and platinum prices through strong discount rate effects, although gold prices fall by less due to higher expected rental income, which is consistent with empirical evidence. I achieve these results by modeling the countercyclical component of gold and platinum service flows in reduced-form. The micro-foundations of this mechanism are an important open question, which I leave for fruitful future research.

# Tables

## Table 1.1: Summary Statistics for Predictors

Table 1.1 gives descriptive statistics for the log GP ratio and other known stock return predictors. Monthly data from 1975 - 2013. $\log GP_t$ is the log GP ratio, computed as the log of the ratio of monthly gold to platinum fixing prices. Monthly prices are the average of daily prices. Prior to April 1990, I use monthly average dealer prices for platinum. Gold fixing prices are from the LBMA, and platinum fixing prices are from the LPPM. Platinum dealer prices are from the USGS. $\log PD_t$ is the log price-dividend ratio for the CRSP value-weighted index. $\log PE_t$ is the cyclically adjusted price-earnings ratio from Robert Shiller's website. $\log PNY_t$ is the net payout yield from Michael Roberts' website, available until December 2010. $ICC_t$ is the implied cost of capital from Li et al. (2013), available from January 1977. $DFSP_t$ is the default spread, calculated as the difference between the yield of Baa and Aaa corporate bonds; the data is from FRED. $TMSP_t$ is the term spread, calculated as the difference in yield between a 10 year constant maturity U.S. government bond and a 3 month constant maturity U.S. treasury bill. The data is from FRED. $INFL_t$ is the growth rate of the consumer price index from the FRED. $CAY_t$ is the consumption-wealth ratio from Lettau and Ludvigson (2001) and the data is from Martin Lettau's website, available until March 2013. The data is quarterly, interpolated to a monthly frequency. $VRP_t$ is the variance premium, calculated as the difference between the squared VIX index and annualized realized volatility over the past month. The VIX data is from the CBOE website, and the high-frequency realized variance is from Hao Zhou's website, available from January 1990. ADF is the augmented Dickey and Fuller (1979) test statistic, and p-val is its p-value. The number of lags in the ADF test is selected based on the Ng and Perron (1995) sequential t-test.

| Variable | Mean | Std. Dev. | AR(1) | ADF | p-val. | Min. | Max. | Corr. $GP_t$ | Start | End |
|---|---|---|---|---|---|---|---|---|---|---|
| $\log GP_t$ | -0.233 | 0.266 | 0.981 | -2.872 | 0.049 | -0.850 | 0.299 | 1.000 | 1975.1 | 2013.12 |
| $\log PD_t$ | 3.608 | 0.447 | 0.994 | -1.238 | 0.657 | 2.764 | 4.510 | -0.588 | 1975.1 | 2013.12 |
| $\log PE_t$ | 2.878 | 0.474 | 0.995 | -1.227 | 0.662 | 1.893 | 3.789 | -0.539 | 1975.1 | 2013.12 |
| $\log PNY_t$ | 2.247 | 0.249 | 0.979 | -2.276 | 0.180 | 1.700 | 3.235 | -0.461 | 1975.1 | 2010.12 |
| $ICC_t$ | 7.445 | 2.694 | 0.949 | -2.759 | 0.064 | -0.040 | 13.850 | 0.339 | 1977.1 | 2013.12 |
| $DFSP_t$ | 1.126 | 0.474 | 0.961 | -4.252 | 0.001 | 0.550 | 3.380 | 0.336 | 1975.1 | 2013.12 |
| $TMSP_t$ | 1.819 | 1.261 | 0.952 | -3.458 | 0.009 | -2.650 | 4.420 | 0.232 | 1975.1 | 2013.12 |
| $INFL_t$ | 0.322 | 0.323 | 0.643 | -4.574 | 0.000 | -1.787 | 1.420 | 0.027 | 1975.1 | 2013.12 |
| $CAY_t$ | 0.003 | 0.018 | 0.995 | -1.798 | 0.381 | -0.035 | 0.034 | 0.258 | 1975.1 | 2013.03 |
| $VRP_t$* | 0.022 | 0.024 | 0.258 | -5.700 | 0.000 | -0.217 | 0.140 | 0.051 | 1990.1 | 2013.12 |

## Table 1.2: U.S. Stock Return Predictability

Table 1.2 shows return predictability regressions for the U.S. equity market, January 1975 to December 2013, 468 monthly observations. The regression is:

$$\frac{12}{h}\sum_{i=1}^{h}\log R_{t+i} - \log R^f_{t+i} = \beta_0 + \beta_1 \log GP_t + \epsilon_{t+h}$$

The left hand variable is the excess log return of the CRSP value-weighted index, annualized by the horizon $h$. The right hand predictor is $\log GP$. Returns are calculated from overlapping monthly data, and t-statistics are based on Newey and West (1987) HAC robust standard errors. The bottom panel estimates a VAR following Hodrick (1992).

| VW Excess Returns | 1m | 3m | 6m | 1y | 2y | 3y | 4y | 5y |
|---|---|---|---|---|---|---|---|---|
| **OLS Regression** | | | | | | | | |
| $\log GP_t$ | 0.237 | 0.246 | 0.260 | 0.243 | 0.202 | 0.161 | 0.145 | 0.129 |
| t-stat. | (2.82) | (3.14) | (2.94) | (2.76) | (2.67) | (3.12) | (4.11) | (4.77) |
| ${R_{adj}}^2$ (%) | 1.21 | 4.23 | 9.55 | 16.57 | 23.60 | 23.57 | 27.80 | 31.66 |
| **VAR Estimation** | | | | | | | | |
| $\log GP_t$ | 0.236 | 0.234 | 0.228 | 0.215 | 0.192 | 0.172 | 0.155 | 0.140 |
| t-stat. | (2.75) | (2.89) | (2.82) | (2.66) | (2.69) | (2.62) | (2.61) | (2.60) |
| ${R_{VAR}}^2$ (%) | 1.27 | 3.46 | 6.34 | 10.89 | 16.43 | 18.99 | 19.84 | 19.73 |

## Table 1.3: Univariate Return Predictability

Table 1.3 shows univariate return predictability regressions for the U.S. equity market, controlling for other known predictors. January 1975 to December 2013, 468 monthly observations. The regression is:

$$\frac{12}{h}\sum_{i=1}^{h}\log R_{t+i} - \log R^f_{t+i} = \beta_0 + \beta_1 X_t + \epsilon_{t+h}$$

The left hand variable is the excess log return of the CRSP value-weighted index return, annualized by the horizon $h$. The right hand predictor is $\log GP$. Returns are calculated from overlapping monthly data, and t-statistics use Newey and West (1987) HAC robust standard errors. ${R_{adj}}^2$ is the adjusted $R^2$ statistic. $X_t$ is a return predictor, including $\log GP$.

| | 1 month horizon | | | 3 month horizon | | | 1 year horizon | | | 5 year horizon | | |
|---|---|---|---|---|---|---|---|---|---|---|---|---|
| | Coef. | t-stat. | $R^2_{adj}$ | Coef. | t-stat. | $R^2_{adj}$ | Coef. | t-stat. | $R^2_{adj}$ | Coef. | t-stat. | $R^2_{adj}$ |
| $\log GP_t$ | 0.237 | 2.82 | 1.21 | 0.246 | 3.14 | 4.23 | 0.243 | 2.76 | 16.57 | 0.129 | 4.77 | 31.66 |
| $\log PD_t$ | -0.067 | -1.19 | 0.10 | -0.066 | -1.39 | 0.68 | -0.065 | -1.30 | 3.15 | -0.064 | -2.94 | 22.12 |
| $\log PE_t$ | -0.045 | -0.86 | -0.05 | -0.045 | -1.00 | 0.24 | -0.050 | -1.04 | 2.01 | -0.049 | -1.99 | 14.70 |
| $\log PNY_t$ | -0.096 | -0.84 | -0.03 | -0.084 | -0.86 | 0.21 | -0.122 | -1.29 | 3.35 | -0.095 | -3.52 | 14.31 |
| $ICC_t$ | 0.019 | 1.98 | 0.74 | 0.014 | 1.75 | 1.22 | 0.015 | 2.33 | 6.08 | 0.013 | 2.86 | 22.35 |
| $DFSP_t$ | 0.013 | 0.18 | -0.20 | 0.016 | 0.25 | -0.16 | 0.035 | 0.88 | 0.90 | 0.032 | 1.07 | 5.23 |
| $TMSP_t$ | 0.015 | 0.77 | -0.08 | 0.013 | 0.77 | 0.04 | 0.025 | 1.80 | 3.77 | 0.017 | 2.15 | 12.41 |
| $INFL_t$ | -0.071 | -0.87 | -0.02 | -0.008 | -0.10 | -0.21 | -0.081 | -1.93 | 2.48 | -0.013 | -0.68 | 0.18 |
| $CAY_t$ | 0.566 | 0.46 | -0.18 | 0.743 | 0.66 | -0.03 | 1.158 | 1.10 | 1.49 | 1.970 | 3.43 | 23.03 |
| $VRP_t$* | 5.011 | 4.32 | 5.12 | 4.370 | 6.27 | 11.12 | 1.265 | 2.60 | 2.95 | -0.487 | 0.99 | 1.90 |

## Table 1.4: Bivariate Return Predictability: Short Horizon

Table 1.4 shows bivariate return predictability regressions for the U.S. equity market for 1 and 3 month horizons, controlling for other known predictors. January 1975 to December 2013, 468 monthly observations. The regression is:

$$\frac{12}{h}\sum_{i=1}^{h}\log R_{t+i} - \log R^f_{t+i} = \beta_0 + \beta_1 \log GP_t + \beta_2 X_t + \epsilon_{t+h}$$

For the monthly frequency, $h = 1$. The left hand variable is the excess logarithmic return of the CRSP value-weighted index return annualized by the horizon $h$. The right hand predictors are $\log GP$ and another return predictor $X_t$. Returns are calculated from overlapping monthly data, and t-statistics use Newey and West (1987) HAC robust standard errors. ${R_{adj}}^2$ is the adjusted $R^2$ statistic.

| | 1 month horizon | | | | | 3 month horizon | | | | |
|---|---|---|---|---|---|---|---|---|---|---|
| | GP Coef. | GP t-stat. | Coef. | t-stat. | $R^2_{adj}$ | GP Coef. | GP t-stat. | Coef. | t-stat. | $R^2_{adj}$ |
| $\log PD_t$ | 0.262 | 2.71 | 0.026 | 0.40 | 1.03 | 0.278 | 3.25 | 0.032 | 0.66 | 4.16 |
| $\log PE_t$ | 0.274 | 2.94 | 0.038 | 0.68 | 1.09 | 0.288 | 3.45 | 0.043 | 0.98 | 4.33 |
| $\log PNY_t$ | 0.232 | 2.44 | 0.019 | 0.16 | 0.80 | 0.250 | 2.65 | 0.040 | 0.39 | 3.54 |
| $ICC_t$ | 0.180 | 1.91 | 0.013 | 1.24 | 1.25 | 0.220 | 2.37 | 0.007 | 0.72 | 4.12 |
| $DFSP_t$ | 0.258 | 2.81 | -0.036 | -0.47 | 1.09 | 0.267 | 3.00 | -0.035 | -0.52 | 4.28 |
| $TMSP_t$ | 0.233 | 2.60 | 0.004 | 0.18 | 1.01 | 0.245 | 2.82 | 0.001 | 0.03 | 4.02 |
| $INFL_t$ | 0.239 | 2.88 | -0.077 | -0.91 | 1.22 | 0.246 | 3.17 | -0.014 | -0.17 | 4.04 |
| $CAY_t$ | 0.241 | 2.64 | -0.345 | -0.26 | 0.99 | 0.250 | 3.03 | -0.201 | -0.18 | 4.03 |
| $VRP_t$* | 0.265 | 2.79 | 4.850 | 4.25 | 6.91 | 0.304 | 3.54 | 4.177 | 6.96 | 18.49 |

## Table 1.5: Bivariate Return Predictability: Long Horizon

Table 1.4 shows bivariate return predictability regressions for the U.S. equity market for 1 and 5 year horizons, controlling for other known predictors. January 1975 to December 2013, 468 monthly observations. The regression is:

$$\frac{12}{h}\sum_{i=1}^{h}\log R_{t+i} - \log R^f_{t+i} = \beta_0 + \beta_1 \log GP_t + \beta_2 X_t + \epsilon_{t+h}$$

For the monthly frequency, $h = 1$. The left hand variable is the excess logarithmic return of the CRSP value-weighted index return annualized by the horizon $h$. The right hand predictors are $\log GP$ and another return predictor $X_t$. Returns are calculated from overlapping monthly data, and t-statistics use Newey and West (1987) HAC robust standard errors. ${R_{adj}}^2$ is the adjusted $R^2$ statistic.

| | 1 year horizon | | | | | 5 year horizon | | | | |
|---|---|---|---|---|---|---|---|---|---|---|
| | GP Coef. | GP t-stat. | Coef. | t-stat. | $R^2_{adj}$ | GP Coef. | GP t-stat. | Coef. | t-stat. | $R^2_{adj}$ |
| $\log PD_t$ | 0.279 | 3.06 | 0.036 | 0.85 | 17.03 | 0.102 | 1.98 | -0.024 | -0.69 | 33.27 |
| $\log PE_t$ | 0.280 | 3.05 | 0.037 | 0.95 | 17.26 | 0.118 | 2.26 | -0.010 | -0.29 | 31.92 |
| $\log PNY_t$ | 0.231 | 2.11 | -0.008 | -0.08 | 14.76 | 0.113 | 2.92 | -0.038 | -0.91 | 33.35 |
| $ICC_t$ | 0.219 | 2.15 | 0.008 | 1.02 | 17.79 | 0.116 | 5.04 | 0.009 | 2.02 | 44.49 |
| $DFSP_t$ | 0.252 | 2.72 | -0.014 | -0.32 | 16.54 | 0.127 | 3.07 | 0.004 | 0.15 | 31.57 |
| $TMSP_t$ | 0.228 | 2.47 | 0.014 | 0.84 | 17.57 | 0.116 | 4.96 | 0.011 | 1.96 | 36.51 |
| $INFL_t$ | 0.248 | 2.92 | -0.089 | -1.94 | 19.65 | 0.131 | 4.97 | -0.020 | -1.40 | 32.57 |
| $CAY_t$ | 0.241 | 2.71 | 0.131 | 0.16 | 16.40 | 0.100 | 4.11 | 1.172 | 1.89 | 38.09 |
| $VRP_t$* | 0.319 | 3.13 | 1.028 | 2.31 | 30.93 | 0.174 | 4.66 | -0.579 | -1.31 | 44.12 |

## Table 1.6: Out-of-Sample Tests

Table 1.6 shows results for out-of-sample testing, using the out-of-sample $R^2$ statistic. Let $T$ be the sample length, and $m$ equal to the size of the initial training window (for expanding regressions) or the size of the training window (for rolling regressions). The Out-of-Sample $R^2$ is given by:

$$R^2_{OS} = 1 - \frac{\sum_{k=1}^{T-m}(r^e_{m+k} - \widehat{r^e}_{m+k})^2}{\sum_{k=1}^{T-m}(r^e_{m+k} - \overline{r^e}_{m+k})^2}$$

For expanding window regressions, the first out-of-sample forecast $\widehat{r^e}_{m+1}$ is based on parameters estimated using observations from 1 to $m$, the second out-of-sample forecast $\widehat{r^e}_{m+2}$ is based on parameters estimated using observations 1 to $m+1$, and so on. For expanding window regressions, the historical average excess return $\overline{r^e}_{t+1}$ is calculated as the average excess return from time 1 to time $t$. For rolling window regressions, the first out-of-sample forecast $\widehat{r^e}_{m+1}$ is based on parameters estimated using observations from 1 to $m$, the second out-of-sample forecast $\widehat{r^e}_{m+2}$ is based on parameters estimated using observations 2 to $m+1$, and so on. For rolling window regressions, the historical average excess return is calculated as the average excess return from over the last $m$ periods, where $m$ is the window length. I consider windows of length 120 months or 180 months to estimate betas, and predict the return in the next month. The In-Sample $R^2$ is the adjusted $R^2$. The p-values are calculated using the adjusted-MSPE statistic of Clark and West (2007) given by:

$$f_{t+1} = (r_{t+1} - \overline{r}_{t+1})^2 - \left[(r_{t+1} - \widehat{r}_{t+1})^2 - (\overline{r}_{t+1} - \widehat{r}_{t+1})^2\right]$$

which is regressed against a constant and the test is a one-sided test.

| | In-Sample (%) | Out-of-Sample (%) | | | | | | | |
|---|---|---|---|---|---|---|---|---|---|
| | | Rolling | | | | Expanding | | | |
| Horizon | | 120m | p-val | 180m | p-val | 120m | p-val | 180m | p-val |
| 1m | 1.21 | -1.46 | 0.315 | 1.44 | 0.018 | 0.88 | 0.033 | 1.35 | 0.028 |
| 3m | 4.23 | 3.95 | 0.010 | 8.28 | 0.002 | 4.59 | 0.009 | 6.18 | 0.008 |
| 6m | 9.55 | 14.39 | 0.002 | 18.29 | 0.001 | 10.89 | 0.010 | 13.39 | 0.010 |
| 1y | 16.57 | 23.02 | 0.009 | 29.37 | 0.007 | 19.31 | 0.021 | 21.35 | 0.024 |
| 2y | 23.60 | 30.98 | 0.044 | 37.28 | 0.030 | 25.64 | 0.041 | 26.02 | 0.045 |
| 3y | 23.57 | 22.08 | 0.069 | 29.30 | 0.027 | 23.56 | 0.033 | 24.51 | 0.029 |
| 4y | 27.80 | 24.03 | 0.047 | 31.39 | 0.010 | 29.36 | 0.008 | 29.69 | 0.004 |
| 5y | 31.66 | 21.09 | 0.057 | 31.88 | 0.010 | 33.60 | 0.010 | 34.37 | 0.003 |

## Table 1.7: International Markets Return Predictability

Table 1.7 shows return predictability regressions for international equity markets, January 1975 to December 2013, 468 monthly observations. The regression is:

$$\frac{12}{h}\sum_{i=1}^{h}\log R_{t+i} - \log R^f_{t+i} = \beta_0 + \beta_1 \log GP_t + \epsilon_{t+h}$$

In Panel A, the left hand variable is the excess log capital gain on the MSCI World Index, annualized by the horizon $h$. The index is calculated in U.S. dollars and the risk-free rate is the U.S. treasury bill rate. Panel B presents the regression results for individual countries, using the respective MSCI country indices denominated in local currency. The risk-free rate for the U.K is the 3-month U.K Treasury rate from FRED. The risk-free rate for Switzerland is the 3-month Swiss franc interbank rate. The risk-free rate for Japan is the interest rate on Japanese Government Treasury bills from FRED. The risk-free rate for Sweden is the 3-month Swedish Treasury rate from FRED from 1982 onwards. Prior to 1982, I use the historical short-term Swedish interest rates from the Sveriges Riksbank website. Newey and West (1987) HAC robust standard errors.

| World Excess Returns | 1m | 3m | 6m | 1y | 3y | 5y |
|---|---|---|---|---|---|---|
| Panel A - World Portfolio | | | | | | |
| $\log GP_t$ | 0.183 | 0.200 | 0.214 | 0.202 | 0.140 | 0.111 |
| t-stat. | (2.15) | (2.22) | (2.02) | (1.88) | (1.94) | (2.52) |
| ${R_{adj}}^2$ (%) | 0.67 | 2.62 | 5.88 | 9.97 | 14.91 | 21.50 |
| Panel B - Individual Countries | | | | | | |
| United Kingdom | | | | | | |
| $\log GP_t$ | 0.266 | 0.239 | 0.232 | 0.213 | 0.156 | 0.105 |
| t-stat. | (3.05) | (3.30) | (3.02) | (2.70) | (2.90) | (4.41) |
| ${R_{adj}}^2$ (%) | 1.19 | 3.46 | 7.34 | 14.15 | 26.00 | 25.52 |
| Switzerland | | | | | | |
| $\log GP_t$ | 0.196 | 0.216 | 0.240 | 0.236 | 0.166 | 0.129 |
| t-stat. | (2.29) | (2.50) | (2.45) | (2.23) | (2.01) | (2.61) |
| ${R_{adj}}^2$ (%) | 0.71 | 2.56 | 6.24 | 11.28 | 16.61 | 19.73 |
| Japan | | | | | | |
| $\log GP_t$ | 0.186 | 0.207 | 0.221 | 0.212 | 0.163 | 0.148 |
| t-stat. | (1.74) | (1.78) | (1.75) | (1.67) | (1.44) | (1.70) |
| ${R_{adj}}^2$ (%) | 0.39 | 1.72 | 3.74 | 6.57 | 10.34 | 15.36 |
| Sweden | | | | | | |
| $\log GP_t$ | 0.391 | 0.427 | 0.427 | 0.367 | 0.190 | 0.132 |
| t-stat. | (2.72) | (3.00) | (2.69) | (2.34) | (1.82) | (2.70) |
| ${R_{adj}}^2$ (%) | 1.54 | 5.08 | 8.88 | 11.51 | 10.62 | 11.96 |

## Table 1.8: Predicting Dividend Growth

Table 1.8 shows dividend growth predictability regressions, January 1975 to December 2013, 468 monthly observations. The regression in Panel A is:

$$\frac{12}{h}\sum_{i=1}^{h}\Delta d_{t+i} = \beta_0 + \beta_1 \log GP_t + \epsilon_{t+h}$$

The regression in Panel B is:

$$\frac{12}{h}\sum_{i=1}^{h}\Delta e_{t+i} = \beta_0 + \beta_1 \log GP_t + \epsilon_{t+h}$$

$\Delta d_t$ is the annualized log dividend growth rate calculated from the CRSP value-weighted index, annualized by the horizon $h$. $\Delta e_t$ is the annualized log earnings growth rate calculated from the earnings data on Robert Shiller's website. The CPI used to deflate dividends is from FRED. Annual dividend growths are calculated from overlapping monthly data, and t-statistics use Newey and West (1987) HAC robust standard errors.

| Panel A - Real Dividend Growth | 1y | 2y | 3y | 4y | 5y |
|---|---|---|---|---|---|
| $\log GP_t$ | 0.018 | 0.027 | 0.028 | 0.023 | 0.014 |
| t-stat. | (0.39) | (0.63) | (0.74) | (0.62) | (0.51) |
| $R^2$ (%) | 0.08 | 0.41 | 0.77 | 0.89 | 0.57 |
| **Panel B - Real Earnings Growth** | **1y** | **2y** | **3y** | **4y** | **5y** |
| $\log GP_t$ | 0.371 | 0.199 | 0.121 | 0.100 | 0.057 |
| t-stat. | (1.31) | (1.06) | (0.93) | (1.12) | (1.04) |
| $R^2$ (%) | 5.08 | 3.61 | 2.79 | 3.48 | 2.78 |

## Table 1.9: Cross-Sectional Implications

Table 1.9 shows the implications of GP risk for the cross-section of stock returns. I first run time-series regressions to estimate betas:

$$R^e_{i,t+1} = c_i + \beta_{i,\Delta gp}\Delta \log GP_{t+1} + \epsilon_{i,t+1}$$

where $R^e_{i,t+1}$ is the excess return for portfolio $i$ and $\Delta \log GP_t + 1 = \log GP_{t+1} - \mathbb{E}_t [\log GP_t]$ is the innovation in GP. The slope coefficient $\beta_{i,\Delta gp}$ represents the portfolio exposure of asset $i$ to GP risk. To estimate the cross-sectional market price of risk associated with GP, I run a cross-sectional regression of time-series average excess returns on the risk exposures:

$$\mathbb{E}\left[R^e_{i,t+1}\right] = \text{cons} + \beta_{i,\Delta gp}\lambda_{\Delta gp} + \upsilon_i.$$

Panel A reports the market price of risk $\lambda$ with Shanken (1992) t-statistic. Panel B shows the results for $\beta_{i,\Delta gp}$ with Newey and West (1987) HAC robust standard errors. The book-to-market and size portfolio returns are from Kenneth French's website. The data is monthly from 1975 - 2013.

| Panel A: Price of Risk | $\lambda_{\Delta gp}$ | t-stat |
|---|---|---|
| Cross-Section | -0.0217 | -6.45 |
| $R^2(\%)$ | 60.64 | |
| **Panel B: Risk Exposures** | $\beta_{\Delta gp}$ | t-stat |
| BM1 | -0.163 | -2.83 |
| BM2 | -0.150 | -2.46 |
| BM3 | -0.137 | -2.93 |
| BM4 | -0.165 | -2.62 |
| BM5 | -0.128 | -2.67 |
| BM6 | -0.170 | -2.79 |
| BM7 | -0.149 | -2.56 |
| BM8 | -0.150 | -2.59 |
| BM9 | -0.177 | -3.12 |
| BM10 | -0.262 | -3.05 |
| SIZE1 | -0.277 | -4.42 |
| SIZE2 | -0.254 | -4.05 |
| SIZE3 | -0.247 | -4.17 |
| SIZE4 | -0.233 | -4.01 |
| SIZE5 | -0.234 | -3.85 |
| SIZE6 | -0.201 | -3.42 |
| SIZE7 | -0.213 | -3.10 |
| SIZE8 | -0.198 | -2.97 |
| SIZE9 | -0.181 | -2.67 |
| SIZE10 | -0.132 | -2.60 |

## Table 1.10: GP and Tail Risk

In Panel A, $SLOPE_t^{n\Delta} = \sigma_t^{OTM,\Delta} - \sigma_t^{ATM}$, where $\sigma_t^{OTM,\Delta}$ is the implied volatility of an out-of-the-money put option with $n\Delta$, where $n = 40, 30, 20$, and $\sigma_t^{ATM}$ is at-the-money implied volatility. Option prices and implied volatilities are from OptionMetrics, January 1996 to August 2013. The regression of the slope of the implied volatility curve for index options against the log GP ratio is:

$$\sigma_t^{OTM,\Delta} - \sigma_t^{ATM} = \beta_0 + \beta_1 \log GP_t + \beta_2 \sigma_t^{ATM} + \epsilon_t$$

The OTM option ranges from deep out-of-the-money ($20\Delta$) to slightly out of the money ($40\Delta$), and the ATM option is defined as a put option with $50\Delta$. The options have just under one month until expiration, and are taken on the last trading day of the month.
In Panel B, the encompassing regression is:

$$SKEW_t^{\mathbb{Q}} = \beta_0 + \beta_1 \log GP_t + \beta_2 \left( VAR_t^{\mathbb{Q}} \times 100 \right) + \epsilon_t$$

where $SKEW_t^{\mathbb{Q}}$ is as defined in Bakshi et al. (2003) and equal to:

$$SKEW_t^{\mathbb{Q}} = \frac{\mathbb{E}_t^{\mathbb{Q}} \left\{ \left( R_{t,t+\tau} - \mathbb{E}_t^{\mathbb{Q}} \left[ R_{t,t+\tau} \right] \right)^3 \right\}}{\mathbb{E}_t^{\mathbb{Q}} \left\{ \left( R_{t,t+\tau} - \mathbb{E}_t^{\mathbb{Q}} \left[ R_{t,t+\tau} \right] \right)^2 \right\}^{3/2}}$$

. For both panels, t-statistics use Newey and West (1987) HAC robust standard errors.

| | $\log GP_t$ | | $\sigma_t^{IV,ATM}$ | | |
|---|---|---|---|---|---|
| Panel A: IV Slope | Coef. | t-stat. | Coef. | t-stat. | ${R_{adj}}^2$ |
| $SLOPE_t^{40\Delta}$ | 0.585 | (2.74) | | | 9.35 |
| | | | 0.056 | (12.54) | 59.26 |
| | 0.332 | (2.97) | 0.053 | (12.91) | 62.13 |
| $SLOPE_t^{30\Delta}$ | 1.402 | (3.10) | | | 11.29 |
| | | | 0.123 | (12.84) | 60.24 |
| | 0.846 | (3.85) | 0.117 | (13.06) | 64.22 |
| $SLOPE_t^{20\Delta}$ | 2.578 | (3.29) | | | 13.05 |
| | | | 0.209 | (11.55) | 59.21 |
| | 1.640 | (4.12) | 0.197 | (11.66) | 64.33 |

| | $\log GP_t$ | | $VAR_t^{\mathbb{Q}} \times 100$ | | |
|---|---|---|---|---|---|
| Panel B: BKM | Coef. | t-stat. | Coef. | t-stat. | ${R_{adj}}^2$ |
| $SKEW_t^{\mathbb{Q}}$ | -0.479 | (-2.84) | | | 5.65 |
| | | | | | -0.06 |
| | | | 0.500 | (3.09) | 5.50 |
| | -0.514 | (-2.71) | | | 5.36 |
| | -0.587 | (-3.58) | 0.615 | (3.62) | 13.97 |
| | -0.560 | (-3.07) | 0.633 | (3.58) | 13.67 |

## Table 1.11: Inventory Financing for Jewellers

Table 1.11 shows the cash ledger for alternative raw materials inventory financing considerations for a hypothetical jeweller deciding between leasing gold or buying gold on credit. $P_{g,t}$ is the spot price of 1 troy oz of gold in period $t$. In both scenarios, the gold ledger is flat. $r_{f,t}$ is the borrowing rate for dollars, $r_{p,t}$ is the futures premium over the spot price, and $\delta_t$ is the gold lease rate. No-arbitrage and no institutional frictions implies $\delta_t = r_{f,t} - r_{p,t}$.

| **Gold Leasing** | **Period** $t$ | **Period** $t+1$ |
|---|---|---|
| Lease gold at $\delta_t$, repay in 1 period | | $-P_{g,t+1} - \delta_t \times P_{g,t}$ |
| Jewellery fabrication & sales | | $P_{g,t+1}$ + Markup |
| Net Cash Flow | 0 | Markup $- \delta_t \times P_{g,t}$ |
| **Buying Gold on Credit** | Period $t$ | Period $t+1$ |
| Borrow money at $r_{f,t}$, repay in 1 period | $+P_{g,t}$ | $-(1 + r_{f,t}) \times P_{g,t}$ |
| Buy spot gold | $-P_{g,t}$ | |
| Jewellery fabrication & sales | | $P_{g,t+1}$ + Markup |
| Short gold futures | | $(1 + r_{p,t}) \times P_{g,t} - P_{g,t+1}$ |
| Net Cash Flow | 0 | Markup $- (r_{f,t} - r_{p,t}) \times P_{g,t}$ |

## Table 1.12: Gold and Platinum Returns

Table 1.12 estimates gold and platinum returns. All estimates are annualized real percentage terms, except for the Sharpe ratios. The data is monthly from 1975 - 2013. The spot prices are calculated from the LBMA fixing price (gold), LPPM fixing price (platinum). Before April 1990, I use USGS dealer prices for platinum. Futures prices are based on the futures contracts with closest to 3 months to maturity from the COMEX division of the CME (formery NYMEX). The interest rate is the U.S. dollar Libor rate (before 1986, the Eurodollar deposit rate). CPI data is from FRED, and the risk-free rate is the 1 month U.S. Treasury bill rate. $\delta$ is the dividend yield (for stocks) or lease rate (convenience yield, for gold and platinum).

Equity, Gold, and Platinum Returns

| Equity | | Gold | | Platinum | |
|---|---|---|---|---|---|
| Variable | Data | Variable | Data | Variable | Data |
| $\mathbb{E}[R^m - R^b]$ | 7.53 | $\mathbb{E}[R^g - R^b]$ | 2.40 | $\mathbb{E}[R^x - R^b]$ | 6.51 |
| $\sigma(R^m)$ | 15.11 | $\sigma(R^g)$ | 16.76 | $\sigma(R^x)$ | 22.18 |
| Sharpe Ratio | 0.50 | Sharpe Ratio | 0.14 | Sharpe Ratio | 0.29 |
| $\mathbb{E}[\delta^m]$ | 2.71 | $\mathbb{E}[\delta^g]$ | 1.00 | $\mathbb{E}[\delta^x]$ | 3.47 |

Table 1.13: Model Parameters

Parameter values below are in annual terms. The third column gives the reference source for the parameter. "Standard" means that the parameter value is in a standard range of values commonly used in the literature. "Data" means that this parameter choice is disciplined by the data. Citations mean this parameter is from the cited paper. "Bounded Parameter" means this parameter is free to be set within bounds imposed by either existence of solutions (for $a_2$ and $b_2$) or by data estimates (for $\mu_\lambda$). "Fixed" is for the intratemporal elasticity of substitution and volatility of $\xi_t$, as discussed in the main text. NSBU (2013) refers to Nakamura et al. (2013) and BU (2008) refers to Barro and Ursua (2008).

| Parameter | Value | Reference |
|---|---|---|
| Relative risk aversion $\gamma$ | 3 | Standard |
| Subject time preference $\delta$ | 0.012 | Standard |
| Mean log consumption growth (normal times) $\bar{g_c}$ | 0.025 | Data |
| Volatility of log consumption growth (normal times) $\sigma_c$ | 0.020 | Data |
| Leverage $\phi$ | 2.6 | Wachter (2013) |
| Mean-reverting target of $\xi_t$ process $\bar{\xi}$ | 0.0355 | Wachter (2013) |
| Rate of mean reversion $\kappa_\lambda$ | 0.25 | Target AR1($\log GP$) |
| Rate of mean reversion $\kappa_\xi$ | 0.095 | Target AR1($p-d$) |
| Volatility of $\lambda_t$ process $\sigma_\lambda$ | 0.183 | Target $\sigma(\log GP)$ |
| Volatility of $\xi_t$ process $\sigma_\xi$ | 0.0861 | Fixed |
| Probability of default given disaster $q$ | 0.40 | Barro (2006), Wachter (2013) |
| Average disaster size $\mathbb{E}_\nu[1-\exp J_t^c]$ | 0.15 | BU (2008), Nakamura et al (2013) |
| Intratemporal elasticity of substitution $\epsilon$ | $1/\phi$ | Fixed |
| Log growth rate in gold and platinum stock $\mu_g$ | 0.0072 | Data |
| Volatility of log gold stock growth $\sigma_g$ | 0.0021 | Data |
| Mean-reverting target of $\log Z_t$ process $\mu_z$ | -3.114 | Data |
| Rate of $\log Z_t$ mean reversion $\theta_z$ | 0.022 | Data |
| Volatility of $\log Z_t$ process $\sigma_z$ | 0.003 | Data |
| Scaling term $a_1 - b_1$ | 6.34 | Match $\mathbb{E}[\log GP]$ |
| Intensity jump mean $\mu_\lambda$ | 0.03 | Bounded Parameter |
| Gold preference loading $a_2$ | 5.73 | Bounded Parameter |
| Platinum preference loading $b_2$ | 1.25 | Bounded Parameter |

## Table 1.14: Simulation Results: Asset Pricing Moments

Table 1.14 shows results from model simulations for stocks, bonds, gold, and platinum. State variables are simulated at a monthly frequency and aggregated to an annual frequency. The population moments are computed from a 1,000,000 year simulation. The model confidence intervals are computed from 100,000 simulations of length equal to the length of the data. Data moments are calculated using monthly observations, from 1975 to 2013 and annualized. Expected returns, yields, and volatilities are given in real percentage terms. $\delta$ represents the dividend yield (for stocks) or lease rate (for gold and platinum). $p - d$ is the log price-dividend ratio.

| | Data | All Simulations | | | | No Disaster Simulations | | |
|---|---|---|---|---|---|---|---|---|
| | Est. | 5% | 50% | 95% | Pop. | 5% | 50% | 95% |
| Stocks and Bonds | | | | | | | | |
| $\mathbb{E}[R_m - R_f]$ | 7.53 | 3.30 | 6.77 | 12.69 | 7.23 | 3.63 | 6.38 | 10.68 |
| $\sigma(R_m)$ | 15.18 | 11.77 | 20.27 | 31.99 | 21.53 | 10.35 | 16.32 | 24.99 |
| $\mathbb{E}[\delta^m]$ | 2.71 | 1.34 | 1.81 | 3.76 | 2.08 | 1.29 | 1.56 | 2.40 |
| $\mathbb{E}[R_f]$ | 1.11 | -1.06 | 2.46 | 3.50 | 2.00 | 1.93 | 3.16 | 3.57 |
| $\sigma(R_f)$ | 1.07 | 0.29 | 1.71 | 6.96 | 3.51 | 0.17 | 0.73 | 2.06 |
| $\sigma(p - d)$ | 0.45 | 0.12 | 0.28 | 0.58 | 0.44 | 0.10 | 0.21 | 0.44 |
| $AR1(p - d)$ | 0.92 | 0.56 | 0.82 | 0.94 | 0.92 | 0.50 | 0.78 | 0.92 |
| Gold | | | | | | | | |
| $\mathbb{E}[R_g - R_f]$ | 2.40 | 0.23 | 2.62 | 5.74 | 2.78 | 0.88 | 2.67 | 4.96 |
| $\sigma(R_g)$ | 16.76 | 6.19 | 10.38 | 20.64 | 12.48 | 5.67 | 7.60 | 11.15 |
| $\mathbb{E}[\delta^g]$ | 1.00 | 0.59 | 0.77 | 1.79 | 0.93 | 0.58 | 0.66 | 0.96 |
| Gold Sharpe | 0.14 | 0.02 | 0.26 | 0.52 | 0.22 | 0.13 | 0.35 | 0.57 |
| Platinum | | | | | | | | |
| $\mathbb{E}[R_x - R_f]$ | 6.51 | 2.37 | 5.34 | 10.22 | 5.70 | 2.70 | 5.05 | 8.55 |
| $\sigma(R_x)$ | 22.18 | 9.37 | 15.28 | 23.98 | 16.27 | 8.40 | 12.30 | 17.49 |
| $\mathbb{E}[\delta^x]$ | 3.47 | 2.33 | 3.01 | 5.45 | 3.34 | 2.26 | 2.63 | 3.67 |
| Platinum Sharpe | 0.29 | 0.15 | 0.37 | 0.60 | 0.35 | 0.24 | 0.42 | 0.62 |
| GP Ratio | | | | | | | | |
| $\mathbb{E}[\log GP]$ | -0.23 | -0.37 | -0.25 | 0.15 | -0.23 | -0.38 | -0.31 | -0.09 |
| $\sigma(\log GP)$ | 0.26 | 0.05 | 0.14 | 0.42 | 0.28 | 0.04 | 0.10 | 0.29 |
| $AR1(\log GP)$ | 0.79 | 0.50 | 0.78 | 0.92 | 0.91 | 0.46 | 0.74 | 0.90 |

## Table 1.15: Simulation Results: Return Predictability

Table 1.15 shows results from model simulations for return predictability. State variables are simulated at a monthly frequency and aggregated to an annual frequency. The population moments are computed from a 1,000,000 year simulation. The model confidence intervals are computed from 100,000 simulations of length equal to the length of the data. I run the regression in the first panel is:

$$\frac{12}{h}\sum_{i=1}^{h}\log R_{t+i} - \log R^f_{t+i} = \beta_0 + \beta_{h,gp}\log GP_t + \epsilon_{t+h}$$

and the period is annual. Excess returns are log equity returns over the log return on the government bond, as in Barro (2006) and Wachter (2013). The regressions I run in the second panel are:

$$\log GP_t = \beta_0 + \beta_Z \log Z_t + \epsilon_t$$

and predicting returns using $\log Z_t$

$$\frac{12}{h}\sum_{i=1}^{h}\log R_{t+i} - \log R^f_{t+i} = \beta_0 + \beta_{h,Z}\log Z_t + \epsilon_{t+h}.$$

In the third panel, I run the return predictability regression controlling for supply effects:

$$\frac{12}{h}\sum_{i=1}^{h}\log R_{t+i} - \log R^f_{t+i} = \beta_0 + \beta_{h,gp\perp Z}\log GP_t + \gamma_{h,Z}\log Z_t + \epsilon_{t+h}.$$

| | Data | All Simulations | | | | No Disaster Simulations | | |
|---|---|---|---|---|---|---|---|---|
| | Est. | 5% | 50% | 95% | Pop. | 5% | 50% | 95% |
| Excess Returns | | | | | | | | |
| $\beta_{1y,gp}$ | 0.243 | -0.007 | 0.334 | 1.023 | 0.132 | 0.095 | 0.444 | 1.148 |
| $\beta_{3y,gp}$ | 0.161 | -0.006 | 0.303 | 0.753 | 0.127 | 0.100 | 0.390 | 0.818 |
| $\beta_{5y,gp}$ | 0.129 | -0.010 | 0.270 | 0.598 | 0.121 | 0.090 | 0.333 | 0.635 |
| $R^2_{1y,gp}$ | 16.57 | 0.28 | 7.69 | 23.01 | 3.07 | 1.22 | 10.36 | 25.94 |
| $R^2_{3y,gp}$ | 23.57 | 0.76 | 20.63 | 49.42 | 8.85 | 3.31 | 26.50 | 53.13 |
| $R^2_{5y,gp}$ | 31.66 | 1.08 | 30.06 | 63.78 | 13.98 | 3.92 | 37.04 | 66.72 |
| Supply Regressions | | | | | | | | |
| $\beta_Z$ | 7.86 | -23.20 | 1.60 | 26.80 | 1.69 | -14.64 | 1.59 | 18.57 |
| $R^2_Z$ | 11.35 | 0.09 | 9.10 | 49.09 | 0.77 | 0.07 | 8.41 | 46.15 |
| $\beta_{1y,Z}$ | -0.55 | -10.19 | 0.01 | 10.26 | 0.00 | -7.49 | 0.03 | 7.63 |
| $R^2_{1y,Z}$ | 0.16 | 0.00 | 1.12 | 9.27 | 0.00 | 0.00 | 1.03 | 8.80 |
| Excess Returns Controlling for Supply | | | | | | | | |
| $\beta_{1y,gp\perp Z}$ | 0.283 | 0.004 | 0.392 | 1.154 | 0.132 | 0.111 | 0.512 | 1.286 |
| $\beta_{3y,gp\perp Z}$ | 0.188 | 0.005 | 0.343 | 0.817 | 0.127 | 0.115 | 0.432 | 0.884 |
| $\beta_{5y,gp\perp Z}$ | 0.145 | 0.002 | 0.293 | 0.634 | 0.121 | 0.105 | 0.358 | 0.668 |
| $R^2_{1y,gp\perp Z}$ | 19.99 | 1.65 | 11.25 | 28.00 | 3.07 | 3.34 | 13.88 | 30.58 |
| $R^2_{3y,gp\perp Z}$ | 28.84 | 4.58 | 28.55 | 56.53 | 8.85 | 9.28 | 33.81 | 59.51 |
| $R^2_{5y,gp\perp Z}$ | 35.65 | 6.88 | 40.28 | 70.68 | 13.98 | 12.80 | 45.87 | 72.80 |

## Table 1.16: Gold and Platinum Supply Dynamics

Table 1.16 shows the per-capita growth rate of the world gold and platinum stock, annual data from 1975 to 2013. The data is calculated from annual world production data. The data for platinum production is from Johnson Matthey. The data for gold production is from the U.S. Geological Survey Minerals Yearbook. The estimate of the initial gold stock is 84,000 tonnes (Thomas and Boyle (1986)), and the estimate of the initial platinum stock (net of autocatalysts) is set to be 4.5% of the initial gold stock. Population growth is proxied by the U.S. annual population growth from the U.S. census. $G_t$ is the per-capita gold stock and $X_t$ is the per-capita platinum stock. Panel A gives descriptive statistics for the log growth rates of per-capita gold and platinum stock. Panel B shows results from augmented Dickey-Fuller tests. Panel C reports the estimate of the cointegrating coefficient $\beta_G$ from a Stock and Watson (1993) Dynamic Least Squares (DLS) regression:

$$\log X_t = \beta_0 + \beta_G \log G_t + \sum_{i=-k}^{k} \gamma_i \Delta \log G_{t-i} + \epsilon_t$$

using Newey and West (1987) HAC robust standard errors. Panel D results of the Johansen (1988) rank test for the VECM:

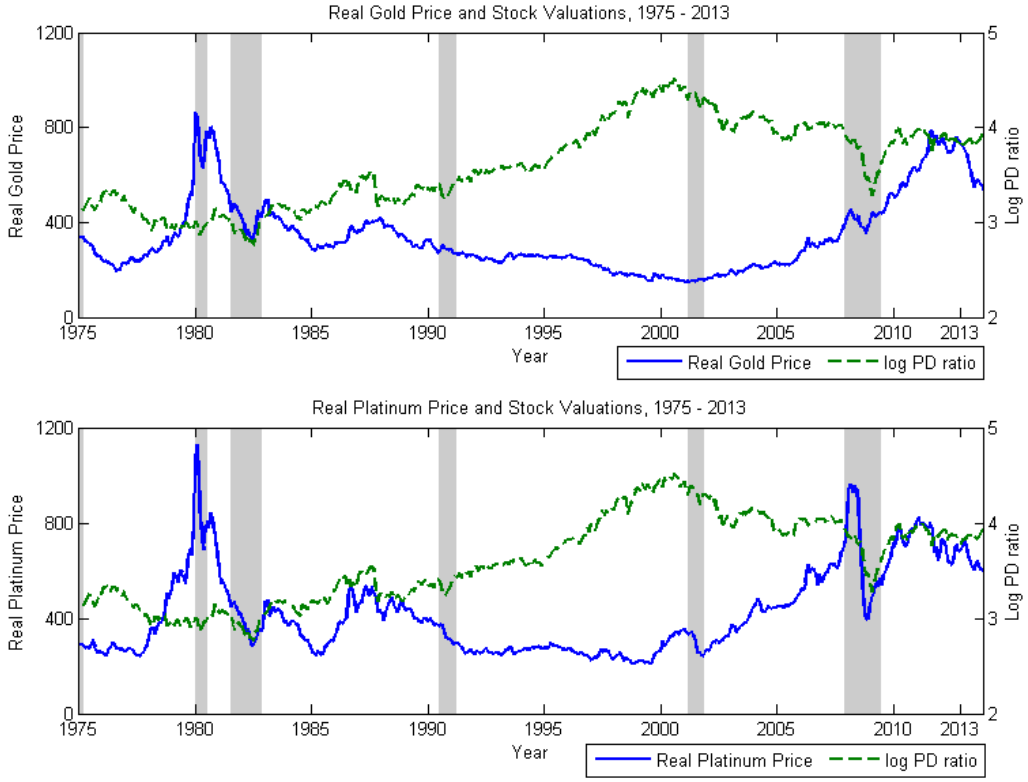
$$\Delta Y_t = \mu + \Pi Y_{t-1} + \sum_{j=1}^{p-1} \Gamma_j \Delta Y_{t-j} + \epsilon_t$$

where $Y_t = \begin{bmatrix} \log X_t & \log G_t \end{bmatrix}'$. The null hypothesis is $H_0 : rank(\Pi) = r$ against the alternative $H_a : rank(\Pi) > r$.

| Panel A - Growth Rates | Mean (%) | Median (%) | Std. Dev. (%) | Min. (%) | Max. (%) |
|---|---|---|---|---|---|
| $\Delta \log G_t$ | 0.72 | 0.76 | 0.21 | 0.21 | 1.00 |
| $\Delta \log X_t$ | 0.71 | 0.68 | 0.29 | 0.24 | 1.28 |
| **Panel B - ADF** | **1 lag** | **2 lags** | **3 lags** | **4 lags** | **5 lags** |
| $\log G_t$ | 1.046 | 1.645 | 1.326 | 1.464 | 1.580 |
| p-val | 0.995 | 0.998 | 0.997 | 0.997 | 0.998 |
| $\log X_t$ | 2.592 | 1.954 | 1.879 | 1.259 | 1.319 |
| p-val | 0.999 | 0.999 | 0.999 | 0.996 | 0.997 |
| $\log Z_t$ | -1.776 | -3.085 | -3.032 | -2.781 | -3.024 |
| p-val | 0.392 | 0.028 | 0.032 | 0.061 | 0.033 |
| **Panel C - DLS** | **Estimate** | **Std. Err.** | **t-stat** | **[95% confidence interval]** | |
| $\beta_G$ (k=1) | 0.990 | 0.034 | 29.39 | 0.92 | 1.06 |
| $\beta_G$ (k=2) | 1.006 | 0.035 | 29.02 | 0.94 | 1.08 |
| $\beta_G$ (k=3) | 1.038 | 0.044 | 23.38 | 0.95 | 1.13 |
| **Panel D - Rank Test** | **Statistic** | **90% CV** | **95% CV** | **p-val** | **$H_0 = r$** |
| 1 lag in VAR | 11.61 | 13.43 | 15.50 | 0.177 | 0 |
| | 0.99 | 2.71 | 3.84 | 0.448 | 1 |
| 2 lags in VAR | 23.28 | 13.43 | 15.50 | 0.004 | 0 |
| | 3.10 | 2.71 | 3.84 | 0.078 | 1 |
| 3 lags in VAR | 26.01 | 13.43 | 15.50 | 0.001 | 0 |
| | 2.53 | 2.71 | 3.84 | 0.112 | 1 |
| 4 lags in VAR | 21.76 | 13.43 | 15.50 | 0.005 | 0 |
| | 0.84 | 2.71 | 3.84 | 0.504 | 1 |

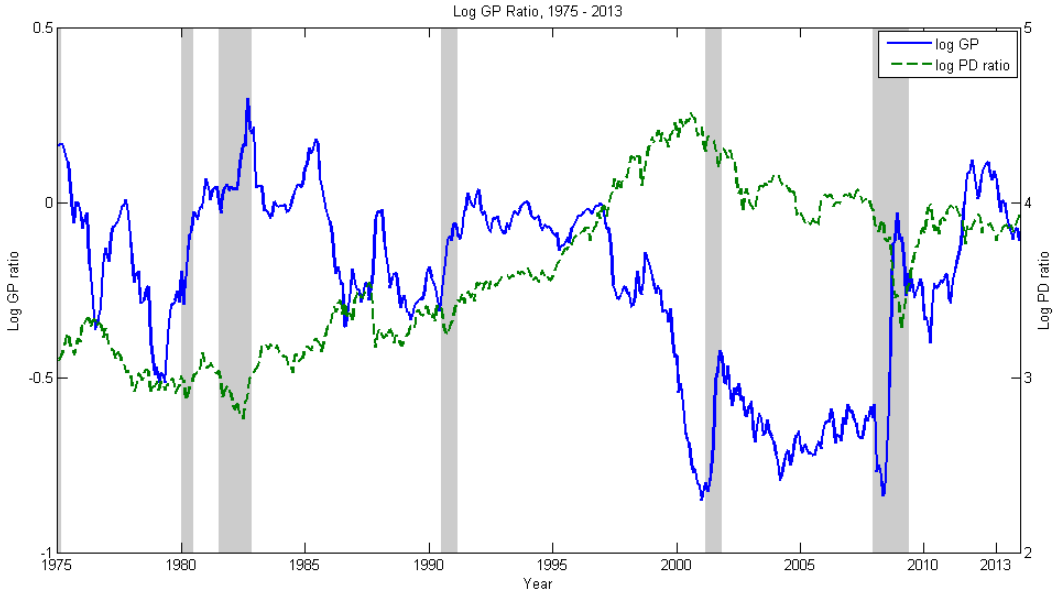
Figures

Figure 1.1: Gold and Platinum Prices



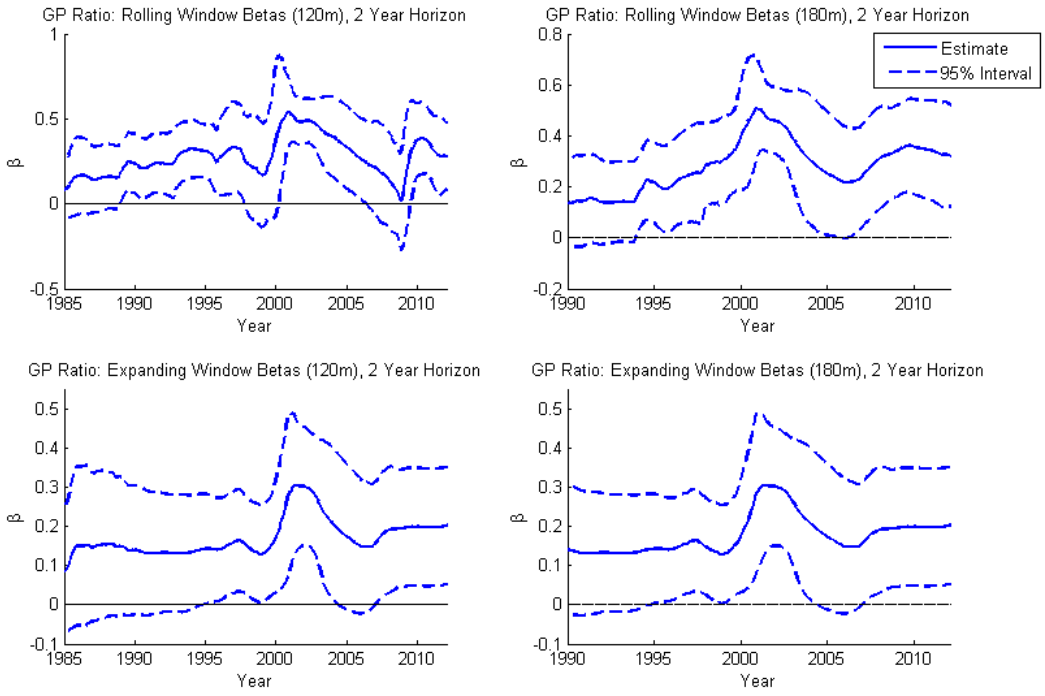

The top panel shows the behavior of real gold prices (solid line) and the log price-dividend ratio on the CRSP value-weighted portfolio (dashed line) from 1975 - 2013. The bottom panel shows real platinum prices (solid line) and the log price-dividend ratio on the CRSP value-weighted portfolio (dashed line) from 1975 - 2013. The shaded grey bars are NBER recessions.

Figure 1.2: Log GP Ratio



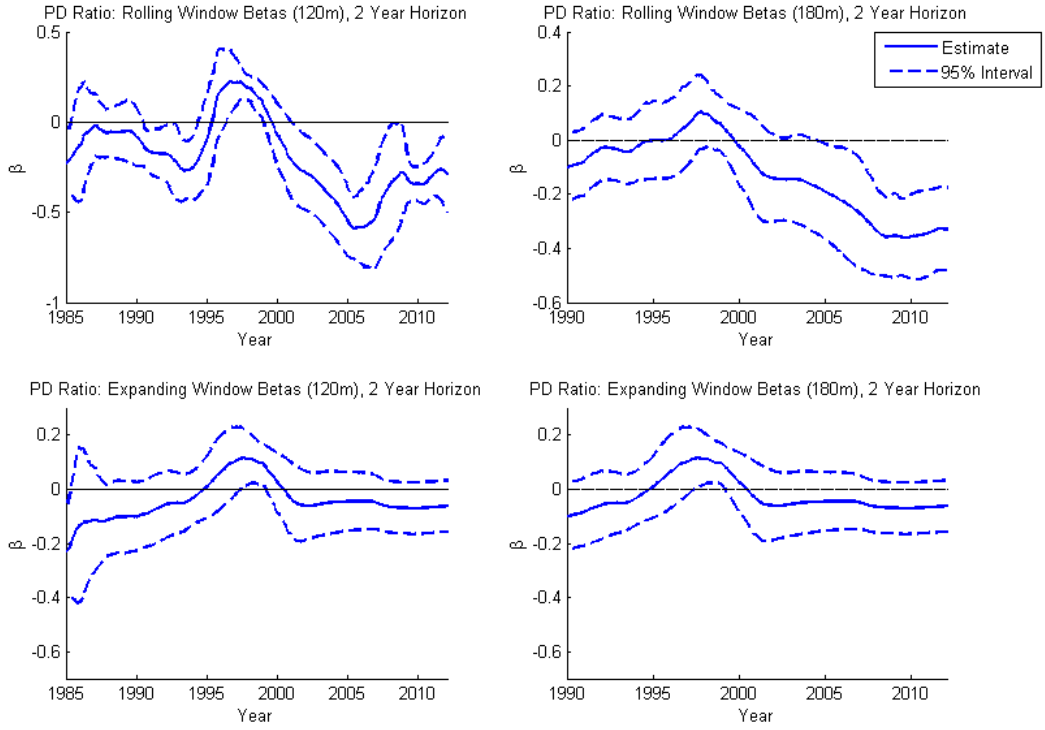

The figure above shows the natural logarithm of the ratio of gold to platinum prices (log GP ratio) from 1975 to 2013. The data is monthly frequency. Gold data is from LBMA, and platinum data is from LPPM and the U.S. Geological Survey. Shaded bars represent NBER recessions.

Figure 1.3: Rolling Regressions - GP ratio



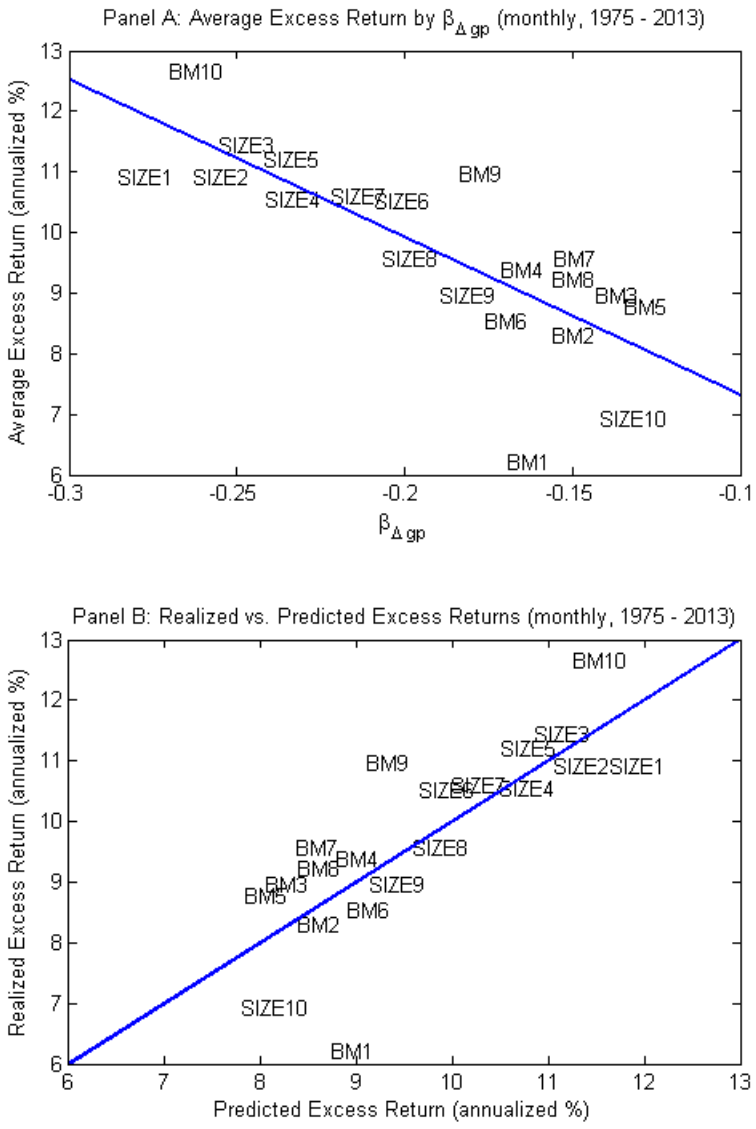

The figure above shows estimated betas and 95% confidence intervals for 2-year ahead predictive regressions of future U.S. stock market excess returns by the log GP ratio. The top left is for the rolling window method with a 120 month window. The top right is for the rolling window method with a 180 month window. The bottom left is for the expanding window method with a 120 month window. The bottom right is for the expanding window method with a 180 month window.

Figure 1.4: Rolling Regressions - PD ratio



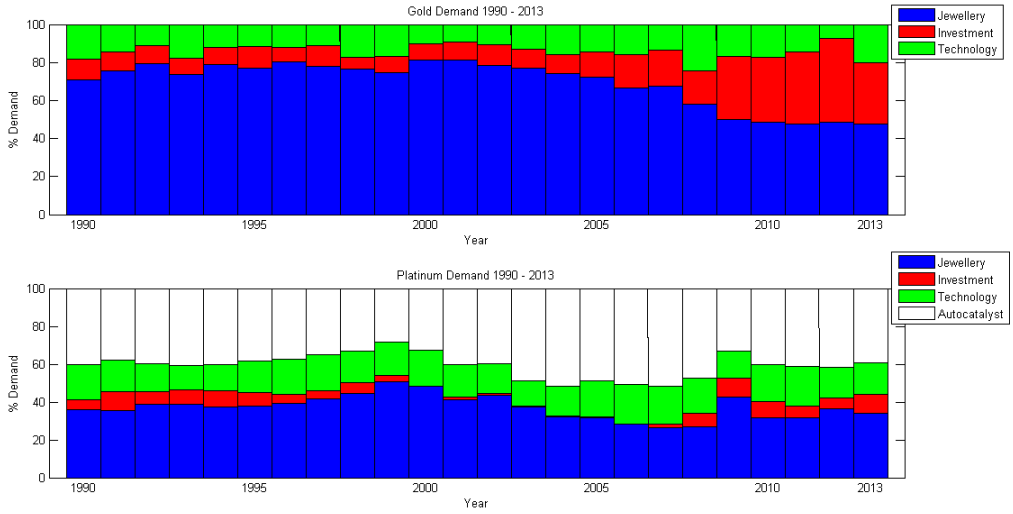

The figure above shows estimated betas and 95% confidence intervals for 2-year ahead predictive regressions of future U.S. stock market excess returns by the log price-dividend (PD) ratio. The top left is for the rolling window method with a 120 month window. The top right is for the rolling window method with a 180 month window. The bottom left is for the expanding window method with a 120 month window. The bottom right is for the expanding window method with a 180 month window.

Figure 1.5: Cross-Sectional Pricing



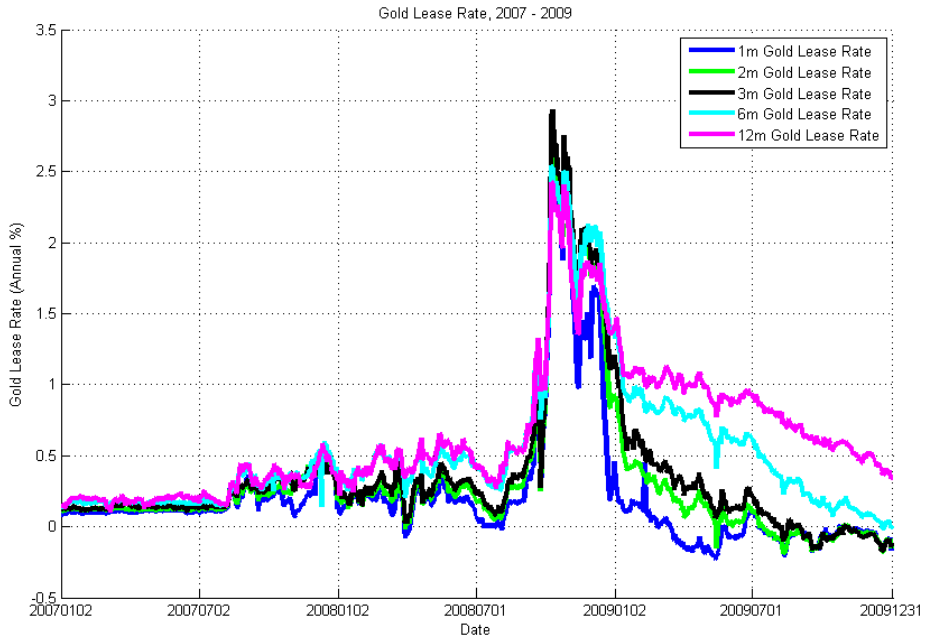

Panel A in the figure above shows the realized average excess returns for book-to-market and size portfolios against the risk exposures (betas). Panel B shows the realized average excess returns against the predicted excess returns. Results are based on the one-factor model with only GP risk. Monthly data 1975 - 2013, annualized percentage excess returns.

Figure 1.6: Gold and Platinum Demand



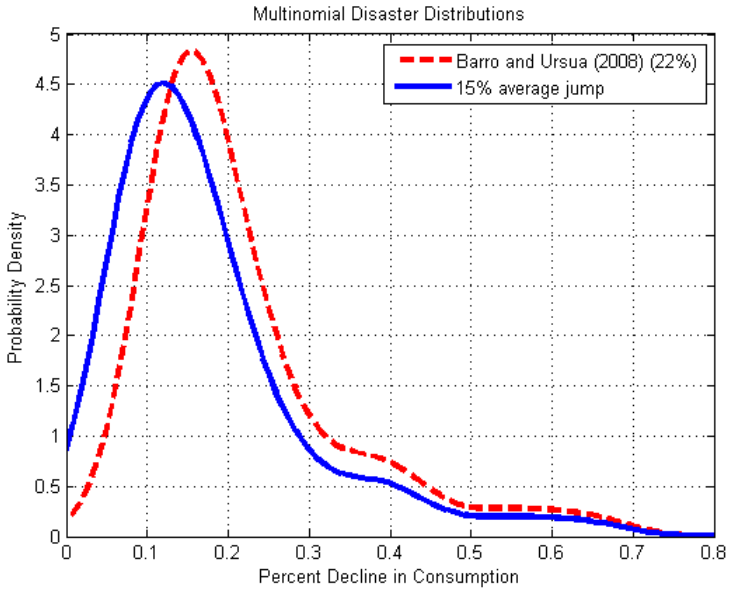

The figure above shows gold and platinum demand for jewellery and investment as a percentage of total demand in a given year, from 1990 - 2013. Gold data is from Thomson Reuters Gold Fields Mineral Services (GFMS) and platinum data is from Johnson Matthey.

Figure 1.7: Gold Lease Rates 2007 - 2009



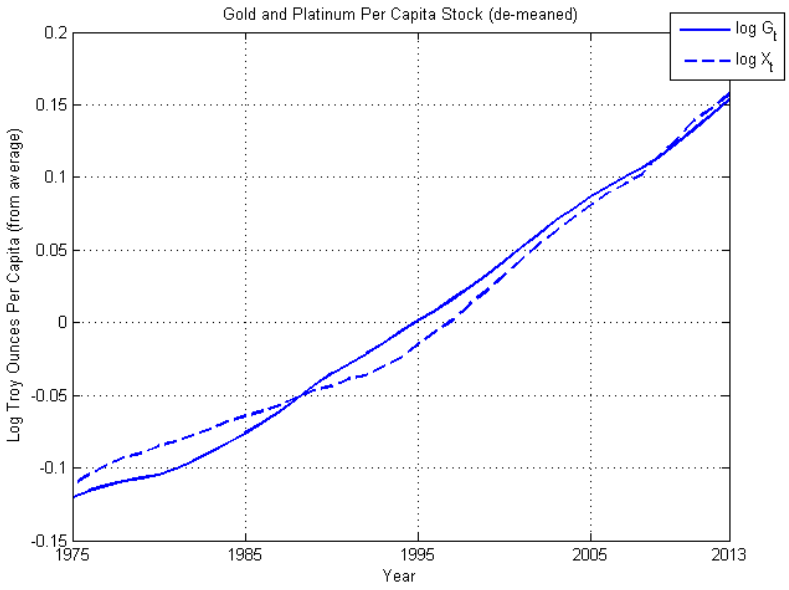

The figure plots the annualized gold lease rates in percentages from 2007 - 2009. The data is computed as the Libor rate minus the Gold Forward Offered Rate (GOFO) as published by the LBMA.

Figure 1.8: Multinomial Disaster Size Distribution



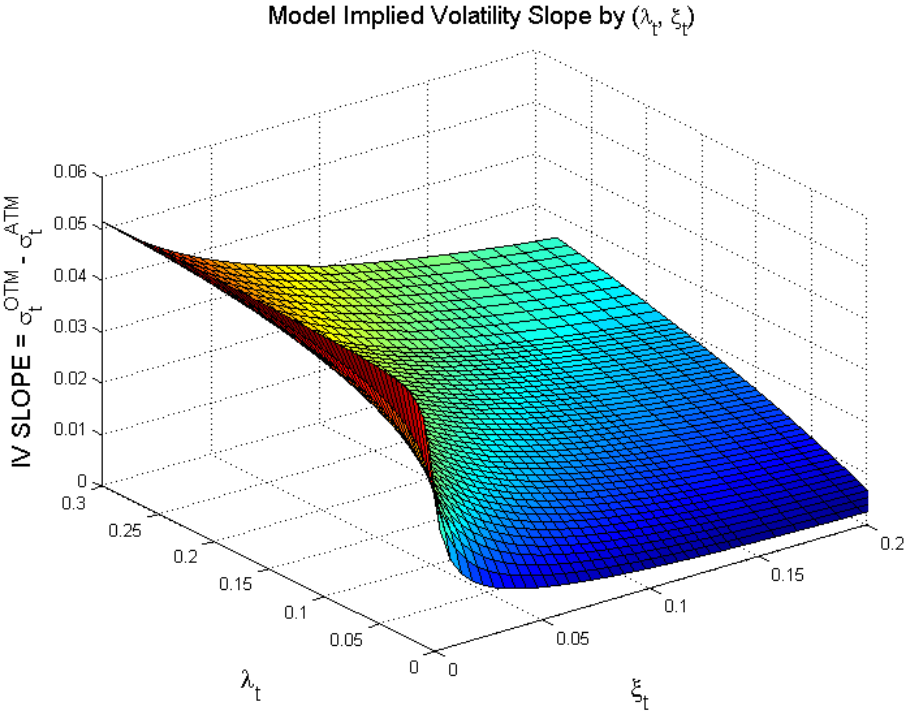

The figure shows the disaster size distribution for the quantity $1 - e^{J_t^c}$ in the model. The red dashed line is the distribution from the Barro and Ursua (2008) calibration, which is also used in Wachter (2013) with an average jump size of 22% using a 10% cutoff to identify disasters. The blue solid line is the distribution used in this paper with an average jump size of 15% using a 6% cutoff to identify disasters. The plots show distributions smoothed by a kernel estimator with a bandwith of 0.05. The model uses draws from the exact multinomial distribution.

Figure 1.9: Per-Capita Gold and Platinum Stock Growth




The figure shows the log per-capita gold and platinum stock (de-meaned) from 1975 to 2013. The data is annual frequency. Gold stock is calculated from world supply data from the U.S. Geological Survey, and platinum stock is calculated from world supply data from Johnson Matthey.

Figure 1.10: Implied Volatility Slope by Disaster Intensities




The figure plots the implied volatility slope, calculated as the difference between a out-of-the-money put option with $\frac{strike}{spot} = 0.95$ and an at-the-money put option with $\frac{strike}{spot} = 1$ with 1 month to maturity, as a function of the state variables $\lambda_t$ and $\xi_t$. Implied volatility measures are calculated by inverting the Black-Scholes formula using the model implied price-dividend ratio and government bond rate.

# CHAPTER 2 : Risk Adjustment and the Temporal Resolution of Uncertainty: Evidence from Options Markets

(with Ivan Shaliastovich)

## 2.1. Introduction

By the Fundamental Theorem of Asset Pricing, the market price of any asset is given by the present value of its payoffs computed under the risk-neutral probability measure (Dybvig and Ross, 1987). These risk-neutral probabilities embed both the objective probabilities of economic states and the risk compensations that investors require for reaching these states. The empirical evidence suggests that the risk-neutral probabilities vary substantially and increase sharply for adverse economic states. Hence, to explain the observed asset prices, economic models need to generate substantial state variations in objective probabilities, the risk compensations, or both. Identifying and quantifying the relative importance of these two channels is one of the central open questions in financial economics, which has significant implications for our understanding of economic risks in financial markets. In this paper, we develop and implement a flexible recursive utility based framework which allows us to separately identify objective probabilities and the risk adjustments using only financial market data, primarily equity index options. We show that risk adjustments, which naturally arise under a preference for early resolution of uncertainty, play a key role in explaining the risk-neutral probabilities observed in the data. Failure to incorporate an effective risk premia channel (e.g., using the expected utility framework) can significantly overestimate the perceived probabilities of bad economic states and underestimate the role of risk adjustments. This, we show, has significant implications for our understanding and interpretation of asset markets.

The theoretical framework of our paper builds on a recent literature which shows that risk-neutral probabilities, observable from option prices, can have substantial information for both physical probabilities and risk adjustments, and under certain conditions can fully

identify the two in the data. Hansen and Scheinkman (2009) establish a general framework to analyze the risk and return relationship, and use an operator approach to theoretically characterize the risk adjustments in Markov state environments. In a more specific application, Ross (2013) shows that physical probabilities and risk adjustments can be recovered together from risk-neutral probabilities when agents' preferences belong to a class of state-independent expected utility models.[37] In this paper, we consider an empirical identification of physical probabilities and risk adjustments in a more general environment which incorporates the Kreps and Porteus (1978) recursive preferences of Epstein and Zin (1989) and Weil (1989). Indeed, when agents' preferences are characterized by state-independent expected utility, the stochastic discount factor only depends on the realized state next period. Then, as shown in Ross (2013), the requirement that implied physical probabilities sum to one provides enough identifying restrictions to separate risk-neutral probabilities into physical probabilities and risk adjustments. However, under recursive preferences, agents also care about current economic conditions which affect the future evolution of wealth. We take this into account and develop a framework to recover physical probabilities and risk adjustments from risk-neutral probabilities in a state-dependent recursive utility environment, given the measurements of the wealth-to-consumption ratio and a preference parameter which captures the preference for the timing of the resolution of uncertainty. By comparing model-implied physical probabilities of the state dynamics to the observed probabilities in the data, we can identify the magnitude of this preference parameter, and thus fully characterize the risk adjustments in the data.

We first use a calibrated economic model to illustrate the role of recursive utility for identifying physical probabilities and risk adjustments. In particular, we consider a Mehra and Prescott (1985) economy with recursive preferences, as in Weil (1989), calibrated to match the dynamics of consumption, dividends, the observed equity premium, and the risk-free rate. We solve our economic model for the equilibrium risk-neutral probabilities and asset

[37]The Appendix in Ross (2013) considers a very special case of recursive preferences where the wealth portfolio is equal to the market and does not pay any dividends. Such a specification can not be supported in equilibrium; further, the full analysis of its implications has not been entertained.

prices, and use them to recover physical probabilities and risk adjustments under alternative assumptions on the preference for the timing of the resolution of uncertainty. We show that unless the econometric framework correctly accounts for a preference for early resolution of uncertainty, we obtain biased estimates of physical probabilities, risk adjustments, and in turn biased estimates of average stock returns.

Intuitively, when the magnitude of the preference for early resolution of uncertainty is not fully accounted for, the recovery methodology trades-off higher probabilities of bad economic states for lower risk compensations in these states. In the limiting case of power utility, it is effectively a manifestation of the failure of the expected utility model to generate sizeable risk premia given the smooth dynamics of the macroeconomy. For example, while the unconditional probability of the lowest consumption growth state is 25% in the benchmark calibration, under the assumption of expected utility the implied physical probability is over 60%. As a consequence of these excessively large probabilities of bad states, the implied population averages of economic variables are significantly biased downward relative to their equilibrium values. For the equity return, its implied average value is less than 1% under expected utility, relative to the 9% calibrated value in the economic model. Hence, the quantitative implications of ignoring the preference for early resolution of uncertainty can be quite significant.

Next, we implement our framework using S&P500 options data. Specifically, we first identify three aggregate economic states based on the past three-month returns on the S&P500 index, with the worst state corresponding to the bottom 25% of the unconditional distribution of market returns and the best state to the top 50% of market returns. In the data, the worst state reflects adverse aggregate economic conditions, with low market returns (-35.51%), low aggregate real consumption growth (0.29%), and a high VIX measure of implied uncertainty (30.29%), relative to high average annualized market returns (22.63%), high real consumption growth (1.79%) and low VIX (20.22%) in the best state. We use S&P500 option prices to estimate the conditional risk-neutral probabilities of future stock

prices in each state following the approach of Figlewski (2008); in particular, the left and right tail of the risk-neutral probability measure are fitted using the Generalized Extreme Value (GEV) distribution. Finally, consistent with the economic model, log wealth-to-consumption ratios are assumed to be proportional to log price-dividend ratios, and the coefficient of proportionality is chosen to match the volatility of the implied wealth-to-consumption ratio in the model to the volatility of the price-dividend ratio in the data.

Using the observed asset prices data, we recover physical probabilities and risk adjustments for different preference parameters. Our empirical findings are in line with the evidence from the economic model. We find that the implied probability of a bad state is quite overstated without a sufficient preference for early resolution of uncertainty. Indeed, when the preferences collapse to expected utility, the implied probability of bad states is about 60%, compared to the its fixed value of 25%. This has a direct implication for the estimated risk compensation across states: the magnitude of the stochastic discount factor going to the bad state is more than two times smaller than going to the good state with expected utility, while the opposite is the case under early resolution of uncertainty. The optimal preference structure which minimizes the distance between the recovered and actual probabilities of the states in the data is very close to the one entertained in our calibrated model, and for risk aversion greater than one, suggests that the inter-temporal elasticity of substitution parameter is above one and that the agent has a a strong preference for early resolution of uncertainty. We verify that these key implications remain similar using various robustness checks on the number and location of the economic states and are quantitatively in line with the existing long-run risks literature (see e.g. Bansal and Yaron, 2004).

Overall, our paper makes a direct contribution to the debate about about the relative importance of objective probabilities relative to risk adjustments to account for the asset prices. Indeed, as shown in Mehra and Prescott (1985), standard expected-utility models fail to capture asset price data because consumption growth is too smooth in the data and the implied risk premia are too low. One approach in the literature is to introduce more

variation in the objective consumption dynamics through rare macroeconomic disasters, either perceived or actual, as in **?** and Barro (2006). Another approach is to enhance the amount of model-implied risk compensation through an alternative preference structure, e.g. the recursive utility of Epstein and Zin (1989). In our paper, we show that the risk adjustments which arise under early resolution play a key role in explaining the variation in risk-neutral probabilities in the data, relative to the objective probabilities. These findings are consistent with Backus, Chernov, and Martin (2011), who show that it is challenging to explain the observed options prices using extreme macroeconomic events under the objective measure.

Our findings are quantitatively consistent with the long-run risks literature which underscores the importance for the inter-temporal elasticity of substitution (IES) to be above one and for a preference for early resolution of uncertainty to explain the key features of asset markets. Bhamra, Kuhn, and Strebulaev (2010) show that a framework with preference for early resolution can generate significant risk premia to account for the observed prices of corporate bonds given low probabilities of defaults in the actual data. Bansal, Dittmar, and Lundblad (2005), Hansen, Heaton, and Li (2008), Bansal, Kiku, Shaliastovich, and Yaron (2013) provide further cross-sectional evidence for these economic channels, and Bansal, Kiku, and Yaron (2007) and Bansal, Kiku, and Yaron (2011), Hasseltoft (2012) consider the empirical evaluation of the model and find support for it in the data.[38]

In our paper, we develop and implement a market-based approach to estimate physical probabilities and risk adjustments separately from the risk-neutral probabilities. Traditionally, the main approach in the literature to identify economic transitions and risk adjustments is to specify parametric models for the physical and risk-neutral dynamics (equivalently, for the physical measure and the risk adjustments); see e.g. Pan (2002), Eraker, Johannes, and Polson (2003), and Eraker (2004). Andersen, Fusari, and Todorov (2012) develop

[38]There is a large literature on the magnitude of the IES. Hansen and Singleton (1982), Attanasio and Weber (1989), Guvenen (2001) and Vissing-Jørgensen and Attanasio (2003) estimate the IES to be in excess of 1. Hall (1988) and Campbell (1999) estimate the IES to be well below one. Bansal and Yaron (2004) argue that the low IES estimates of Hall and Campbell are based on a model without time varying volatility.

a parametric estimation approach for state recovery based on a panel of options prices. Alternatively, Jackwerth (2000), Aït-Sahalia and Lo (2000), and Bliss and Panigirtzoglou (2004) use non-parametric methods to estimate the physical and risk-neutral distributions from the data to derive implications for equilibrium risk adjustments and preference parameters of the representative agent. **?** uses a similar method involving transformations of the payoff function adapted from Bakshi et al. (2003) to obtain model-free estimates of the moments of the risk-neutral distribution from option prices. Carr and Yu (2012) develop the recovery framework in a bounded diffusion context based on restricting the dynamics of the numeraire portfolio.

In the context of equilibrium models, Shaliastovich (2010) and Eraker and Shaliastovich (2008) consider implications of the recursive utility structure and jumps in economic uncertainty for the equilibrium pricing of options, while Liu, Pan, and Wang (2005) and Gabaix (2012) consider the role of rare jumps in explaining out-of-the money option prices. Bekaert and Engstrom (2009) introduce a non-Gaussian bad environment - good environment specification of the consumption dynamics and show it can realistically capture risk-neutral distributions of equity returns in the data. Bollerslev et al. (2009) and Drechsler and Yaron (2011) discuss the implications for the equilibrium risk-neutral variance dynamics and its relation to the asset markets.

Our paper proceeds as follows. In Section 1 we describe our theoretical framework. In Section 2 we set up our model economy, and use model calibrations to highlight the biases that arise without properly considering the preference for early resolution of uncertainty. Section 3 is devoted to our empirical analysis, and Section 4 concludes the paper.

## 2.2. Theoretical Framework

### *2.2.1. Setup of the Economy*

We consider a Markovian regime-switching economy with $N$ discrete states. As is typical in the literature, we specify our economy to be stationary in growth rates, in line with classic

models such as Mehra and Prescott (1985) and Weil (1989) where consumption growth rates and dividend growth rates are stationary and follow a regime-switching process. Under the stationarity of growth rates, the levels of consumption and asset prices follow random walk processes and are unbounded, which is economically appealing; Dubynskiy and Goldstein (2013) discuss further drawbacks for the recovery methods under stationary consumption levels.

Let $m_{i,j}$ denote the stochastic discount factor (SDF), which in standard structural models corresponds to the intertemporal marginal rate of substitution of the representative agent between state $i$ today and state $j$ next period. According to no-arbitrage pricing conditions, the value of an asset is given by the expectation of its payoff times the SDF. In particular, let $q_{i,j}$ denote the price of an Arrow-Debreu security, which is a state-contingent claim in state $i$ that pays one unit of consumption in state $j$ next period. From the pricing equation, the Arrow-Debreu price is given by the product of the SDF and the physical transition probability $p_{i,j}$:

$$q_{i,j} = m_{i,j}\, p_{i,j}. \tag{2.1}$$

By construction, the sum of the Arrow-Debreu prices given the current state is equal to the price of a one-period risk-free bond in this state, and the Arrow-Debreu price scaled by the price of a risk-free bond, $q_{i,j} / \sum_j q_{i,j}$, corresponds to the risk-neutral probability of going from state $i$ today to state $j$ next period. Following Breeden and Litzenberger (1978), we can estimate the risk-neutral probabilities in a model-free manner from the cross-section of options prices (see Section 3.2). Hence, the Arrow-Debreu prices can be directly identified using market data on option and bond prices alone.

Arrow-Debreu prices incorporate information on both physical probabilities and risk adjustments, as shown in (2.1). To identify the physical probabilities and risk adjustments separately, a standard approach in the literature is to use separate parametric or non-parametric models for the physical and risk-neutral dynamics (equivalently, physical dynamics and the

SDF), which are then estimated in the data. Hansen and Scheinkman (2009) establish an alternative framework and use operator approach to theoretically characterize the risk adjustments in Markov state environments. Specifically, as shown in Ross (2013), under restrictions on the class of underlying preferences in the economy, the physical transition probabilities and the SDF can be recovered directly from the Arrow-Debreu prices. We highlight the main steps of this approach below, and then extend the market-based identification framework to the case of recursive utility.

*2.2.2. Expected Utility Structure*

Consider a specification where the representative agent has state- and time-independent expected utility. In this case, the marginal utility between consumption in states $i$ and $j$ depends only on the state the economy is transitioning to and not on the current state. For example, in the classic model of Mehra and Prescott (1985), the agent has expected power utility over future consumption. In this case, the SDF is independent of the current state, and is given by:

$$m_{i,j} = m_j = \delta \lambda_j^{-\gamma}, \tag{2.2}$$

where $\delta$ is the agent's time preference parameter, $\lambda_j$ is the consumption growth rate in state $j$, and $\gamma$ is the coefficient of relative risk aversion. With power utility, the SDF is directly related to the endowment dynamics and the preference parameters of the agent. Our subsequent analysis is more general and is not limited to power utility preferences. In fact, it is valid under any specification of time- and state-independent expected utility, and only relies on the identification assumption that the SDF does not depend on the current state.

Let $Q$ denote the matrix of Arrow-Debreu prices, and $P$ the matrix of physical transition probabilities. Let $M$ be the diagonal matrix with the stochastic discount factors $m_j$ on the

diagonal. Then, we can rewrite (2.1) in matrix form:

$$Q = PM. \tag{2.3}$$

or more explicitly,

$$\begin{bmatrix} q_{1,1} & \dots & q_{1,N} \\ \vdots & \ddots & \dots \\ q_{N,1} & \dots & q_{N,N} \end{bmatrix} = \begin{bmatrix} m_1\, p_{1,1} & \dots & m_N\, p_{1,N} \\ \vdots & \ddots & \dots \\ m_1\, p_{N,1} & \dots & m_N\, p_{N,N} \end{bmatrix}. \tag{2.4}$$

With $N$ discrete states, this is a system of $N^2$ equations in $N^2 + N$ unknowns. As pointed out in Ross (2013), to identify the physical probabilities and the SDF separately one can use the further restriction that conditional physical probabilities must sum to 1:

$$P\,\mathbf{1} = \mathbf{1} \tag{2.5}$$

This gives us $N$ additional restrictions, which allows us to identify uniquely both the physical transition probabilities $P$ and the marginal utilities $M$. Indeed, using the restrictions for the physical probability matrix in (2.3) and (2.5), we obtain:

$$P\,\mathbf{1} = QM^{-1}\mathbf{1} = \mathbf{1} \implies M^{-1}\mathbf{1} = Q^{-1}\mathbf{1}. \tag{2.6}$$

The last condition completely characterizes $M$ since it is a diagonal matrix:

$$M = \left[diag\left(Q^{-1}\mathbf{1}\right)\right]^{-1}. \tag{2.7}$$

We can further recover the implied physical probability transition matrix:

$$P = Q\ diag\left(Q^{-1}\mathbf{1}\right). \tag{2.8}$$

By construction, each row of $P$ sums to one, so subject to the requirement that the recovered $P$ has all positive elements, it represents a valid matrix of conditional probabilities. The implied physical probabilities and the SDF satisfy the pricing equation (2.3), and further, by construction, this decomposition is unique.

Notably, in the analysis above, the physical probabilities and risk adjustments are identified using market data from Arrow-Debreu prices alone, and do not rely on the exact functional form of the SDF, the specification of the endowment process or the agent's preference parameters. Hence, the framework can be implemented empirically using only market data on options and bonds, which are arguably better measured relative to macroeconomic variables. In the next section we consider a recursive utility setup which no longer admits a state-independent SDF. We show that we can maintain a convenient market-based approach for extracting the physical probabilities and the SDF separately from the data, given measurements of the wealth-to-consumption ratio and the preference parameter capturing the preference for the timing of resolution of uncertainty.

*2.2.3. Recursive Preferences Structure*

For the Kreps and Porteus (1978) recursive utility of Epstein and Zin (1989) and Weil (1989), the life-time utility of the agent $V_t$ satisfies,

$$V_t = \left[ (1-\delta) C_t^{1-\frac{1}{\psi}} + \delta \left( \mathbb{E}_t \left[ V_{t+1}^{1-\gamma} \right] \right)^{\frac{1-\frac{1}{\psi}}{1-\gamma}} \right]^{\frac{1}{1-\frac{1}{\psi}}} \tag{2.9}$$

where where $\delta$ is the time discount factor, $\gamma \geq 0$ is the risk aversion parameter, and $\psi \geq 0$ is the intertemporal elasticity of substitution (IES). For ease of notation, the parameter $\theta$ is defined as $\theta \equiv \frac{1-\gamma}{1-\frac{1}{\psi}}$. Note that when $\theta = 1$, that is, $\gamma = 1/\psi$, the above recursive preferences collapse to the standard case of expected utility. As is pointed out in Epstein and Zin (1989), in this case the agent is indifferent to the timing of the resolution of uncertainty of the consumption path. When risk aversion exceeds the reciprocal of IES, the agent prefers early resolution of uncertainty of consumption path, otherwise, the agent has a preference

for late resolution of uncertainty.

Note that when the risk aversion coefficient $\gamma$ is bigger than one, the sign and the magnitude of $\theta$ have a direct economic implication for the preference structure of the agent. In this case, $\theta < 0$ implies that the inter-temporal elasticity of substitution is bigger than one, and further, the agent has a preference for early resolution of uncertainty ($\psi < 1/\gamma$). The IES goes below one when $\theta$ is positive, and when $\theta > 1$ the agent has a preference for late resolution of uncertainty; the limiting case of $\theta = 1$ nests standard expected utility. The relative values of the IES and the preference for the temporal resolution of uncertainty embedded in $\theta$ have significant implications for the equilibrium choices of the agent and for asset prices. For example, the long-run risks literature argues for the inter-temporal elasticity of substitution above one and for a preference for early resolution of uncertainty to explain the key features of asset markets, which suggests $\theta$ is negative (see Bansal and Yaron (2004)). An alternative interpretation of $\theta$ arises in the robust control and model uncertainty literature (see e.g. Hansen and Sargent (2006)), where the parameter $\theta < 0$ captures the agent's aversion to model mis-specification.

As shown in Epstein and Zin (1989), the real stochastic discount factor implied by these preferences is given by

$$m_{i,j} = \delta^{\theta} \lambda_j^{-\frac{\theta}{\psi}} R_{c,i,j}^{\theta-1} \tag{2.10}$$

where $\lambda_j$ is the growth rate of aggregate consumption and $R_{c,i,j}$ is the return on the asset which delivers aggregate consumption as its dividends each time period (the wealth portfolio). This return is different from the observed return on the market portfolio as the levels of market dividends and consumption are not equal: aggregate consumption is much larger than aggregate dividends. Let us decompose the consumption return into its cash flow growth rate, $\lambda_j$, and the change in price-dividend ratio on the wealth portfolio $PC$ between states $i$ and $j$ :

$$R_{c,i,j} = \lambda_j \frac{PC_j + 1}{PC_i}. \tag{2.11}$$

Substitute the return decomposition above into the recursive utility SDF in (2.10) to obtain:

$$m_{i,j} = \left[\delta^{\theta}\lambda_j^{-\gamma}(PC_j + 1)^{\theta-1}\right] PC_i^{1-\theta}. \tag{2.12}$$

Therefore, in the case of the recursive preferences, the SDF depends on both current and future economic states. Unlike the power utility case where agents care just about the next-period consumption shock, with recursive preferences they are concerned about the endogenous dynamics of their wealth, so that the economic variables which affect their wealth now determine the marginal rates of substitution between the periods. When $\theta = 1$, the preferences collapse to standard expected utility in (2.2), and the SDF depends only on the next-period consumption growth rate $\lambda_j$.

Note that the Epstein-Zin SDF can be decomposed multiplicatively into a component which depends only on the next-period state, $\tilde{m}_j$, and the component which depends only on the current state through $PC_i^{1-\theta}$ :

$$m_{i,j} = \tilde{m}_j {PC_i}^{1-\theta}. \tag{2.13}$$

The SDF component which involves the next-period state, $\tilde{m}_j$, in general depends on the endowment dynamics and all the preference parameters. However, if we can characterize the current-state component ${PC_i}^{1-\theta}$, we can extend the recovery approach and identify the physical probabilities and the SDF directly using market data alone. Indeed, let us modify the matrix equation for the Arrow-Debreu prices (2.3) in the following way:

$$Q = PC(\theta)P\tilde{M}, \tag{2.14}$$

where $PC(\theta)$ and $\tilde{M}$ are the diagonal matrices of $PC_i^{1-\theta}$ and $\tilde{m}_i$, respectively. Then, from the condition that physical probabilities sum to one, we obtain:

$$P\mathbf{1} = PC(\theta)^{-1}Q\tilde{M}^{-1}\mathbf{1} = \mathbf{1} \implies \tilde{M}^{-1}\mathbf{1} = Q^{-1}PC(\theta)\mathbf{1}, \tag{2.15}$$

which uniquely characterizes the SDF and the physical probabilities in terms of the Arrow-Debreu prices and the wealth-to-consumption ratios:

$$\begin{aligned} \tilde{M} &= \left[ diag\left(Q^{-1}PC(\theta)\mathbf{1}\right)\right]^{-1}, \\ P &= PC(\theta)^{-1}Q\ diag\left(Q^{-1}PC(\theta)\mathbf{1}\right). \end{aligned} \tag{2.16}$$

Hence, given measurements of the price-dividend ratio on the wealth portfolio and the preference parameter $\theta$, we can still recover the SDF and physical probabilities separately from Arrow-Debreu prices alone, without using macroeconomic data.

An important parameter which affects the the implied physical probabilities and the SDF is the preference parameter $\theta$. When $\theta = 1$, the preferences collapse a to power utility, in which case the expressions for the implied physical probabilities and the stochastic discount factor reduce to their expected utility counterparts in (2.7)-(3.1). For $\theta \neq 1$, the identification of the probabilities and risk adjustments now depends on the magnitudes of the price-dividend ratios in each state. If the price-dividend ratios vary across the states, which we show is empirically relevant, this impacts the inference on the implied physical probabilities and the risk adjustments.

One of the qualitative implications of the recursive structure has to do with the impact of the risk-free rates on the physical probabilities. Ross (2013) demonstrates that if the risk-free rates are constant across the states, the physical probabilities reduce to the risk-neutral probabilities in the world with expected utility. This, however, is no longer the case with recursive preferences. Indeed, even when the risk-free rates are constant, the variation in price-dividend ratios is still going to impact the measurements of physical probabilities and create a wedge between them and the risk-neutral ones. With recursive preferences, the physical probabilities equal to the risk-neutral probabilities only when both the risk-free rates and wealth-to-consumption ratios do not vary across the states.

In the next section, we use a calibrated model to illustrate the quantitative importance of the recursive preferences to identify the physical probabilities and risk adjustments.

Then, we implement this framework in the data and argue that the market data supports a specification with preference for early resolution of uncertainty ($\psi > 1, \theta < 1$).

## 2.3. Economic Model

### *2.3.1. Economy Dynamics*

Consider a discrete-time, infinite horizon endowment economy. The representative agent has recursive utility over future consumption described in (2.9), which allows for a preference for the timing of resolution of uncertainty. The endowment growth $\lambda_j$ follows a time-homogeneous Markov process with a transition matrix $P$, as in Weil (1989) and Mehra and Prescott (1985). Notably, in our model economic growth rates and asset returns are stationary, while the levels of consumption and asset prices follow a random walk process.

Given our specification of the endowment dynamics and the agent's preferences, we can characterize the equilibrium stochastic discount factor $m_{ij}$ in (2.10), and use a standard Euler equation,

$$E_i m_{ij} R_{ij} = 1, \tag{2.17}$$

to compute the equilibrium prices of financial assets, such as Arrow-Debreu assets, the wealth portfolio and the risk-free bonds. To evaluate model implications for the equity market, following Abel (1999) we model equity as a leveraged claim on aggregate consumption and specify its dividend dynamics in the following way:

$$\lambda_j^d = 1 + \mu_d + \phi(\lambda_j - \mu - 1), \tag{2.18}$$

where $\mu_d$ is the mean dividend growth and $\phi > 1$ is the dividend leverage parameter. For simplicity, we abstract from separate equity-specific shocks, so that aggregate dividend growth is perfectly correlated with consumption growth.

*2.3.2. Model Calibration*

The model is calibrated on a quarterly frequency to match the key features of U.S. real consumption growth and financial asset market data from 1929 to 2010. Specifically, we first start with an AR(1) dynamics of consumption growth on a quarterly frequency:

$$\lambda_{t+1} = 1 + \mu + \rho(\lambda_t - \mu - 1) + \sigma\epsilon_{t+1}, \tag{2.19}$$

where $\epsilon_{t+1}$ is a Normal shock. We calibrate the mean $\mu$, persistence $\rho$ and volatility $\sigma$ parameters so that a long sample of consumption growth simulated from the above specification and time-aggregated to an annual frequency targets the mean, volatility and persistence of the annual consumption growth in the data. To calibrate the dividend process, we set the average dividend growth to be the same as the mean consumption growth and fix the dividend leverage parameter to $\phi = 3$, similar to other studies. We then discretize the AR(1) dynamics of consumption growth into 3 states using the Tauchen and Hussey (1991) quadrature approach.[39]

The calibrated parameter values are presented in Table 2.1, and the key moments for consumption growth in the model and in the data are shown in Table 2.2. In the data, the consumption corresponds to annual real expenditures on non-durables and services from the BEA tables; in the model, the population moments are computed from a long simulation of a Markov process for quarterly consumption growth, time-aggregated to an annual horizon. As shown in Table 2.2, our calibration ensures that the model matches very closely the key properties of consumption growth in the data. Both in the model and in the data, the average consumption growth is 1.90%, and its persistence is about 0.50 on annual frequency. The volatility of consumption growth is 2.50% in the model, close to 2.20% in the data.

We calibrate the preference parameters $\delta, \gamma$ and $\psi$ to match the key moments of the financial asset market data. The subjective discount factor $\delta$ is set at 0.988, annualized. The risk

[39]Our main results are virtually unchanged using alternative number of economic states, or specifying the model on an annual frequency.

aversion is calibrated to $\gamma = 25$ and the inter-temporal elasticity of substitution parameter $\psi$ is equal to 2.[40] As shown in Table 2.2, our model delivers an average market price-dividend ratio of about 60, a market risk premium of 7% and the risk-free rate of 1.5%, which is consistent with the data. The implied volatility of the market return is about 10% in the model, which is lower than its typical estimates in the data of $15 - 20\%$. Recall that for simplicity, our model does not entertain equity-specific shocks in the dividend dynamics, which can help match the volatility of dividends and returns in the data.

Notably, as in the long-run risks literature, we focus on the case when both the inter-temporal elasticity of substitution and risk aversion parameters are above one, so that $\theta = (1-\gamma)/(1-1/\psi) = -48$ is below zero. As discussed in Bansal and Yaron (2004), this parameter configuration is economically appealing for several reasons. First, keeping $\psi > 1$ ensures that the substitution effect dominates wealth effect, which implies that the equilibrium equity valuations fall in bad times of low economic growth. For example, in our calibration, the market price-dividend ratio is 62.77 in the best consumption growth state (state 1), and it drops to 57.12 in the worst state (state 3), and similar for the wealth portfolio. Further, under such a parameter configuration the agent has preference for early resolution of uncertainty ($\psi > 1/\gamma$), and thus dislikes negative shocks to expected consumption or increases in aggregate volatility. These model predictions are directly supported in the data, and motivate a preference parameter configuration where $\theta < 0$.[41]

### *2.3.3. Model Implications: Consumption States*

To highlight the main intuition for our results, we first consider the identification of the physical probabilities and the risk adjustments in the benchmark model specification using the consumption states. In the next section, we present the analysis under the market

[40] Our risk aversion coefficient is higher than in Bansal and Yaron (2004) who use a monthly calibration of the model. Note that our model is specified on a quarterly frequency and does not entertain a separate expected growth component. See Bansal et al. (2011) for further discussion of time-aggregation issues in measurements of the preference parameters.

[41] See Bansal et al. (2005) and Hansen et al. (2008) for the cross-sectional evidence, Bansal et al. (2007) and Bansal et al. (2011) for the empirical evaluation of the model.

return-based states.

We use our model to compute the equilibrium Arrow-Debreu prices $Q$ between the three consumption states, and further decompose them into the physical probability $P$ and the risk adjustment through the stochastic discount factor $M$:

$$\underbrace{\begin{bmatrix} 0.47 & 0.50 & 0.03 \\ 0.07 & 0.61 & 0.33 \\ 0.01 & 0.17 & 0.82 \end{bmatrix}}_{Q} = \underbrace{\begin{bmatrix} 0.68 & 0.31 & 0.01 \\ 0.17 & 0.67 & 0.17 \\ 0.01 & 0.31 & 0.68 \end{bmatrix}}_{P} \times \underbrace{\begin{bmatrix} 0.69 & 1.60 & 3.44 \\ 0.39 & 0.91 & 1.96 \\ 0.24 & 0.56 & 1.21 \end{bmatrix}}_{M}, \tag{2.20}$$

where $\times$ indicates element-by-element multiplication. Note that the SDF depends both on the current and future states. The SDF value is highest going from the best state 1 which has the highest consumption growth to the worst state 3 in which consumption growth is lowest, and smallest values of SDF obtain when we transition from the worst 3 to the best consumption state 1. The Arrow-Debreu prices incorporate the effects of both the risk adjustment and the physical transition probabilities. In this case, because going from the best to the worst state is very unlikely, the Arrow-Debreu price between states 1 and 3 is relatively inexpensive. On the other hand, consumption in the worst state next period is very valuable given current worst state, which can be attributed both to a relatively high risk compensation for remaining in state 3 and the persistence of the Markov chain.

The Arrow-Debreu price decomposition above is based on the equilibrium solution of the model given the full calibration of the endowment dynamics and preference parameters. Let us now consider the case when the researcher only has access to the model-generated data on Arrow-Debreu prices and the wealth-to-consumption ratios and tries to identify the physical probabilities of economic states and the implied risk compensation. Following our discussion in Section 2.2.3, the identified values are based on the preference parameter $\theta$, so we consider a range of possible values for $\theta$ and use the conditions in (2.16) to recover the implied physical probabilities and SDF at the entertained values of $\theta$.

In the top two panels of Figure 2.1 we show the implied unconditional probability of being in the bad state, and the implied value for the SDF which corresponds to transitioning from the good to the bad state relative to remaining in the good state (i.e. $\frac{m_{1,3}}{m_{1,1}}$). When $\theta$ is equal to its calibrated value of $-48$, the unconditional probability of remaining in the bad state and the relative value of the SDF in the bad state are equal to their equilibrium values of 25% and 5.02, respectively. When the candidate value of $\theta$ is smaller in the absolute value relative to its calibrated value, the recovery is based on the understated magnitude of the risk compensation, leading to biased estimates of physical probabilities and risk adjustments. Specifically, the Figure shows that the probabilities of bad events are biased upwards, while the risk adjustments of bad events are biased downwards.

Consider, for example, the decomposition of the Arrow-Debreu prices into the implied physical probabilities and implied SDF under the assumption of expected utility, so that $\theta$ is set to 1:

$$\underbrace{\begin{bmatrix} 0.47 & 0.50 & 0.03 \\ 0.07 & 0.61 & 0.33 \\ 0.01 & 0.17 & 0.82 \end{bmatrix}}_{Q} = \underbrace{\begin{bmatrix} 0.47 & 0.50 & 0.03 \\ 0.07 & 0.61 & 0.33 \\ 0.01 & 0.17 & 0.82 \end{bmatrix}}_{P(\theta=1)} \times \underbrace{\begin{bmatrix} 0.99 & 1.00 & 1.00 \\ 0.99 & 1.00 & 1.00 \\ 0.99 & 1.00 & 1.00 \end{bmatrix}}_{M(\theta=1)}. \tag{2.21}$$

Under the expected utility, the SDF does not depend on the current state, so all the rows in the matrix $M(\theta = 1)$ are identical. Further, as can be seen from the above equation, under the expected utility there is barely any variation in the SDF across future states, compared to the benchmark SDF in (2.20). This is consistent with the broad literature which finds that time- and state-independent expected utility models can not generate enough volatility of the SDF to account for the asset market prices; for example, the risk-free rate and the equity premium puzzles imply that standard expected utility models cannot simultaneously explain the levels of the risk-free rate and equity risk premium given the actual dynamics of consumption growth in the data (see e.g., Mehra and Prescott (1985)). Following the fact that under the expected utility assumption there is very little action coming from the risk

compensation, virtually all the difference in the Arrow-Debreu prices is now attributed to the difference in the implied physical probabilities. To account for the observed asset market features, the recovery framework needs to twist the physical dynamics of the endowment, and in particular, it needs to put more weight on the likelihood of bad events to generate expensive Arrow-Debreu prices of going into bad states. Because of that, under expected utility the unconditional probability of bad states is significantly biased upwards and equal to 62% versus its true value of 25%, as shown in Figure 2.1. Similar biases for the stochastic discount factor and physical probabilities arise at alternative values of $\theta$ which are less than the calibrated one in absolute value. These values of preference parameter underestimate the volatility of the SDF and the amount of risk compensation it can generate, and thus lead to higher implied probabilities of bad events, as shown in Figure 2.1.

The mis-specification of the economic dynamics has important implications for the implied moments of macroeconomic variables and financial market variables, as shown in the last column of Table 2.2 where we document the moments of consumption growth and stock returns computed using the recovered physical probabilities under the expected utility assumption. As evident in the Table, attributing more likelihood to bad events leads to a significant downward bias of the measured average consumption growth rate and the mean returns. For example, while the mean consumption growth rate is calibrated to 1.9%, using the recovered physical probabilities under the expected utility framework, the implied mean is -0.8%. Similarly, as shown in the bottom panel of Figure 2.1, the average return on the market implied by the physical probabilities under expected utility is less than 1%, relative to its calibrated value of 8.9%.

To formally evaluate the mis-specification of the economic dynamics, we compute the Kullback-Leibler (KL) divergence between the implied and calibrated physical probabilities. This is the standard measure of the fit of distributions, and is calculated as the distance between the true distribution and a candidate distribution; smaller values of KL

divergence imply a better fit. It is given by the following equation:

$$KL(P||P_\theta) = \sum_i \pi_i \left[ \sum_j p_{i,j} \log \left( \frac{p_{i,j}}{{p^\theta}_{i,j}} \right) \right], \tag{2.22}$$

where $\pi_i$ are the unconditional probabilities of being in each state, $p_{i,j}$ are the calibrated physical transition probabilities, and ${p^\theta}_{i,j}$ are the recovered transition probabilities for a particular candidate value of $\theta$. We plot the value of KL divergence for candidate values of $\theta$ on the bottom right panel of Figure 2.1. The criterion function is minimized at zero (no deviation from the true distribution) at the true value of $\theta$, and significantly rises for alternative values of $\theta$.

Overall, our findings suggest that while analysis based on the expected utility framework suggests a high probability of bad events (e.g., disasters), these results have to be interpreted with caution and might just indicate a mis-specification of the underlying preference structure of the agent. Indeed, when the preference structure allows for a preference for the timing of the resolution of uncertainty, the burden of explaining the cross-section of Arrow-Debreu prices falls less on the physical probabilities, and the differences in Arrow-Debreu prices are attributed more to variations in risk compensation across states.

### *2.3.4. Model Implications: Market Return States*

In the previous Section, we considered consumption growth regimes to identify the aggregate economic states and decompose the Arrow-Debreu prices for these states into physical probabilities and risk adjustments. However, using the observed index option prices in the data, we can only identify the risk-neutral distributions and the Arrow-Debreu prices corresponding to the states of aggregate market returns. In this section, we use our model simulation to analyze the implications of the alternative identification of the states based on market returns for the recovery of preference parameters and physical probabilities.

Specifically, given the assumption of the three underlying consumption states, the distribu-

tion of realized market returns consists of 9 possible return values, $R_{ij}$, for $i, j = 1, 2, 3$. We identify the best return state corresponding to the top third of the distribution of market returns; in our calibration, the best return state includes all three returns going to the best consumption state, and the return from the worst to the middle consumption state. The worst return state corresponds to the bottom third of the distribution of returns, and includes all three returns going to the worst consumption state. Finally, the middle return state includes the intermediate values of returns. We compute the physical probability and the risk-neutral probability matrices, average values of wealth-to-consumption ratio and the risk-free rate, and the Arrow-Debreu prices for these return-based states. We can then use the recovery methodology to decompose the Arrow-Debreu prices into the implied physical probabilities and risk adjustments, as in the previous section.

Overall, as the return-based states generally co-move with the consumption-based states, return-based state recovery leads to similar conclusions for the importance of the recursive utility to correctly identify the physical probabilities and the risk adjustments. However, the added noise and averaging out in the measurement of the aggregate states through returns tends to decrease the persistence of the states and diminishes the variability of the Arrow-Debreu prices across states, which brings down the magnitude of the implied risk compensation relative to the calibrated model.

Indeed, the transition matrix for the return states is given by,

$$P^r = \begin{bmatrix} 0.56 & 0.16 & 0.28 \\ 0.17 & 0.67 & 0.17 \\ 0.28 & 0.16 & 0.56 \end{bmatrix}, \tag{2.23}$$

and features a much lower persistence of 0.27 than 0.50 for the consumption states. The Arrow-Debreu prices for the return states also have less variation across the states, relative

to the consumption states, and are given by:

$$Q^r = \begin{bmatrix} 0.18 & 0.43 & 0.38 \\ 0.07 & 0.61 & 0.33 \\ 0.17 & 0.02 & 0.81 \end{bmatrix}. \tag{2.24}$$

Next we decompose the Arrow-Debreu prices for the return states into the implied physical probabilities and the risk adjustments for alternative values of $\theta$, following the approach outlined in the previous section, and report the results in Figure 2.2. Similar to our previous findings, positive values of $\theta$ significantly overestimate the probability of bad states, underestimate the risk adjustments for bad states, and lead to substantial negative biases for the average returns. The preference structure which brings these magnitudes close to the calibrated values relies on strongly negative values of $\theta$. Notably, as the return identification is not based on the consumption states, the parameter which minimizes the KL distance between the calibrated and implied physical probabilities no longer results in a perfect match of the two probabilities, and is different from the calibrated value. Indeed, due to less variation in Arrow-Debreu prices for the return based states, the implied magnitude of risk compensation is lower, and the value of $\theta$ of -12 is smaller than for the consumption-based states. Thus, while the return-based identification lead to similar conclusions for the importance of $\theta < 0$ to account for the asset prices, the added noise and averaging out in the measurement of aggregate states by using return-based states bias down the overall magnitude of the preference parameter $\theta$.

## 2.4. Empirical Analysis

### *2.4.1. Data*

We use the OptionMetrics database to obtain daily closing prices for exchange-traded S&P500 index options on the CBOE from 1996 to 2011. On each trading day, there are an average 840 put and call options contracts written on the S&P500 index and differing

with respect to the expiration date and the strike price; however, a significant number of the contracts is subject to liquidity concerns, such as zero trading volume and large bid-ask spreads. To mitigate possible microstructure issues, we follow Figlewski (2008) to apply standard data filters and exclude contracts with zero trading volume, quotes with best bid below \$0.50, and very deep-in-the-money options. Further, our benchmark analysis is conducted on a quarterly frequency, where we use options with 3 months to maturity and track their prices on the expiration dates of the contract. We focus on the quarterly frequency for several reasons. First, the main liquidity in the options markets lies in the primary quarterly expiration cycle: the main hedging instruments for the options are the S&P500 futures which feature quarterly expirations, so the majority of S&P500 options also trade on the primary quarterly cycle with expirations in March, June, September, and December of each year. Second, in our empirical analysis, aggregate economic states are identified from the distribution of market returns, and focusing on a relatively lower quarterly frequency helps to reduce non-systematic noise in prices. Finally, using quarterly frequency in the data allows us to directly relate our findings to the economic model in Section 2. We have verified our findings are robust to using a monthly data horizon, and we report the results in Section 4. In addition to options prices, we use data on interest rates which correspond to the 3 month U.S. Treasury rate, and the returns and price-dividend ratio on the S&P500 index.

As the options data and the implied risk-neutral probabilities are based on the S&P500 index, we use the distribution of the capital gains on the index, $r^m_{t+1} = \log \frac{S_{t+1}}{S_t}$, to identify the aggregates state of the economy. As evident from the histogram of the capital gains in Figure 2.3, the return distribution is fat-tailed and negatively skewed. Indeed, the skewness of capital gains over the 1996-2011 period is -0.73, and its kurtosis is 4.62 on a quarterly frequency, which is higher than for a normal distribution. Large negative moves in quarterly returns are likely to contain important information about the aggregate economy and are in general more important from the perspective of a risk-averse investor relative to large positive shocks in returns. Motivated by such considerations, we identify 3 economic states,

good, medium, and bad, where the good state corresponds to the upper 50% percentile of the return distribution, the bad state represents the lowest 25% of the returns, and the medium state is in between. The median return in each bin identifies the level of return in each of the economic states, and is given by:

$$\begin{bmatrix} r_1^m & r_2^m & r_3^m \end{bmatrix} = \begin{bmatrix} 22.63\% & 0.19\% & -35.51\% \end{bmatrix}, \tag{2.25}$$

annualized. The estimated transition matrix for the states in the data is equal to,

$$P = \begin{bmatrix} 0.53 & 0.34 & 0.13 \\ 0.47 & 0.20 & 0.33 \\ 0.44 & 0.13 & 0.44 \end{bmatrix}. \tag{2.26}$$

Because the good state corresponds to the upper 50% of the return distribution, there is a considerable probability of remaining in the good state (53%) or transitioning to the medium one (34%). The overall persistence of the aggregate state implied by the transition matrix is low and matches the persistence of returns in the data. The persistence of returns in the data is 0.13, while the persistence of the estimated Markov chain above is 0.23. The persistence of the return states is also very similar to the value in the economic model of 0.27, as discussed in Section 2.4.

We verify that the states identified by the returns on the index contain meaningful information about the aggregate economy, and we report the average values of the key economic variables, such as real consumption growth, VIX and asset prices, in Table 2.3. As shown in the Table, there is a significant difference between the average economic variables in the two extreme states. The median real consumption growth is 1.79%, annualized, in the good state, whereas it is a much lower 0.29% in the bad state. The VIX index, which measures uncertainty about the market, is 20.22 in the good state, relative to a much higher value of 30.29 in the bad state, and the price-dividend ratio for the index increases from about 50 to 55 going from the bad state to the good state. Looking at Figure 2.4, we see that the implied

volatility curves for each state generated from options data are increasing across the range of moneyness as the aggregate state worsens. In particular, at-the-money implied volatility increases from 18% in the good state to 26% in the bad state. Overall, our economic states meaningfully capture the real growth and uncertainty prospects in the aggregate economy.

*2.4.2. Estimation of the Risk-Neutral Distribution*

Theoretically, the entire risk-neutral probability distribution can be extracted directly using a continuum of options contracts, as shown in Breeden and Litzenberger (1978). Let $\tilde{P}(x) = \int_{-\infty}^{x} \tilde{p}(z)\, dz$ denote the risk neutral cumulative distribution function, $K$ be the strike price, and $r$ the risk-free interest rate. Then, given the current value $S$ of the underlying, the price of a European call option expiring at time $T$ is given by:

$$\begin{aligned} C(K;S,r) &= PV\{\mathbb{E}^{Q}[\max(S_T - K, 0)]\} \\ &= \int_{K}^{\infty} e^{-rT}(S_T - K)\, \tilde{p}(S_T)\, dS_T. \end{aligned}$$

Differentiating the call price with respect to strike price allows us to relate the risk-neutral distribution to the prices of call options:

$$\frac{\partial C}{\partial K} = e^{-rT}\left[-(K-K)\,\tilde{p}(K) + \int_{K}^{\infty} -\tilde{p}(S_T)\, dS_T\right] = -e^{-rT}\left[1 - \tilde{P}(K)\right], \tag{2.27}$$

so that the risk-neutral probability is determined by the second derivative of the price of call options:

$$\tilde{p}(K) = e^{rT}\frac{\partial^2 C}{\partial K^2}. \tag{2.28}$$

In practice, we do not observe the entire continuum of options prices, and we do not observe very deep in- and out-of-the money contracts to capture the tails of the distribution. To address the first issue, we interpolate the data to fill in the quotes between listed strikes. Following Shimko (1993), we first transform option prices into Black and Scholes (1973a) implied volatilities and interpolate the implied volatility surface, and then transform the

interpolated curve back to find a theoretical profile of call option prices by strike. As we are interested in the conditional probabilities of being in the good, medium and bad aggregate states, we use quarterly data to calculate the average volatility surface in each of the three states and then compute the implied risk-neutral distributions conditional on each state. To deal with the unavailability of deep in- and out-of-the money contracts, we follow the steps in Figlewski (2008) and approximate the tails of the risk-neutral density by a Generalized Extreme Value (GEV) distribution. The tail parameters of the distribution are chosen to match the curvature of the risk-neutral probability density function at two extreme points, along with the requirement that the tail probabilities in both the observed risk-neutral distribution and the GEV tail distributions must equal. A similar approach is also pursued by Vilkov and Xiao (2012). All of the details for the estimation of the risk-neutral distribution are provided in the Appendix.

Figure 2.5 shows the estimated risk-neutral distributions for each of the aggregate states, together with the GEV adjustments of the right and left tails of the distribution. The bottom panel of Table 2.3 summarizes the conditional moments of the risk-neutral distribution. Going from good to bad state, risk-neutral volatility increases from 18.92% to 25.27% ; the risk-neutral 3rd central moment becomes about two times more negative, and the 4th moment of the distribution increases twofold as well. Overall, good states are characterized by relatively lower volatility and a relatively lower left tail, which is consistent with our findings on the behavior of VIX and asset prices in the previous section. Our evidence is also consistent with Chang, Christoffersen, and Jacobs (2013) who find that risk-neutral skewness correlates negatively with market returns.

### *2.4.3. Data Implications for Probabilities and Risk Adjustments*

Using our estimates of the risk-neutral distributions, we can compute the conditional risk-neutral probabilities between the three aggregate states implied by the options prices. The risk-neutral probabilities, adjusted by the risk-free rates, allow us to calculate the matrix

of Arrow-Debreu prices which is specified below:

$$Q = \begin{bmatrix} 0.47 & 0.14 & 0.38 \\ 0.48 & 0.13 & 0.38 \\ 0.50 & 0.09 & 0.41 \end{bmatrix}. \tag{2.29}$$

The Arrow-Debreu prices appear relatively low for bad states: for example, the Arrow-Debreu price of going from the good state to the bad state is 0.38, relative to 0.50 for going from the bad state to the good state. Ex-ante, it is not clear whether the difference in these prices is attributable to the difference in physical transition probabilities (as suggested by our estimate in (2.26), since going from good to bad is less likely than going from bad to good), or by the difference in the magnitudes of risk compensation between the states. To separate the Arrow-Debreu prices into the implied physical probabilities and the risk-adjustment, we implement our market-based recovery methodology outlined in Section 1, allowing for recursive state-dependent utility and a preference for the timing of the resolution of uncertainty.

The recovery of the probabilities and the SDF relies on measurements of the wealth-to-consumption ratio and the preference parameter $\theta$. The wealth-to-consumption ratio is not directly observed in the data. Consistent with our economic model, we assume that the log wealth-to-consumption ratio is proportional to the log price-dividend ratio, $\log PC \approx \alpha + \beta \log PD$, and set the scale parameter $\beta$ to match the volatility of the wealth-to-consumption ratio in the model relative to the volatility of the price-dividend ratio over the long historical sample in U.S.. Given our economic model, the implied estimate of $\beta \approx 1\%$, which is consistent with empirical findings in Lustig, Nieuwerburgh, and Verdelhan (2012) that the wealth-to-consumption ratio is less volatile than the price-dividend ratio. We examine the robustness of our results to the scale parameter $\beta$ in Section 4.

Given these measurements of the wealth-to-consumption ratio, we entertain a range of possible preference parameters $\theta$ and identify the implied physical probability distribution

and the SDF for each of the values of this parameter. For all the values of $\theta$, the top panels of Figure 2.6 depict the implied unconditional probability of being in the bad state and the implied value of the SDF for transitions from the good to the bad state, while the bottom panels show the implied average market returns, computed under the implied physical probabilities, and the Kullback-Leibler divergence criterion between the recovered transition probabilities and estimated transition probabilities in the data. The actual estimates for the probability of being in bad states, the value of the SDF and the implied average market returns in the case of early resolution of uncertainty and the expected utility are provided in Table 2.4.

Our empirical findings based on the options data are consistent with the output of the economic model in Section 2. As shown in the bottom right panel of Figure 2.6, the implied and actual physical probabilities of the states are best matched when $\theta$ is sufficiently negative, and the the Kullback-Leibler divergence criterion is minimized at $\theta = -11.28$, which if very close to the model value based on the return-based states.

Measurements of $\theta$ have important implications for the recovery of the physical probabilities and the risk adjustments. As in the economic model, the recovered probability of the bad state is significantly higher when the preference parameter $\theta$ is positive or not sufficiently negative. Indeed, for $\theta = 1$ the preferences collapse to expected utility and the implied probability of the bad state is about 60%, compared to its set value of 25%. When $\theta$ equals its Kullback-Leibler optimal value of -11.28, the representative agent has a strong preference for early resolution of uncertainty, and the recovered physical probability of the bad state is close to the actual value in the data. The economic channel which accounts for the upward bias in the recovered probability of bad events is the one highlighted in the economic model: expected utility features very little risk adjustment across states, so physical probabilities have to bear all the burden of explaining the cross-section of Arrow-Debreu prices in the data.

To further illustrate the importance of the recursive utility structure, consider the decom-

position of the Arrow-Debreu prices into implied physical probabilities and the SDF in the expected utility framework and one that features a preference for early resolution of uncertainty,

$$\begin{aligned}\underbrace{\begin{bmatrix}0.47 & 0.14 & 0.38\\ 0.48 & 0.13 & 0.38\\ 0.50 & 0.09 & 0.41\end{bmatrix}}_{Q} &= \underbrace{\begin{bmatrix}0.55 & 0.18 & 0.28\\ 0.55 & 0.17 & 0.28\\ 0.58 & 0.12 & 0.30\end{bmatrix}}_{P(\theta=-11.28)} \times \underbrace{\begin{bmatrix}0.87 & 0.78 & 1.36\\ 0.87 & 0.79 & 1.37\\ 0.86 & 0.77 & 1.35\end{bmatrix}}_{SDF(\theta=-11.28)} \\ &= \underbrace{\begin{bmatrix}0.26 & 0.17 & 0.57\\ 0.27 & 0.16 & 0.57\\ 0.28 & 0.11 & 0.61\end{bmatrix}}_{P(\theta=1)} \times \underbrace{\begin{bmatrix}1.79 & 0.82 & 0.67\\ 1.79 & 0.82 & 0.67\\ 1.79 & 0.82 & 0.67\end{bmatrix}}_{SDF(\theta=1)}.\end{aligned} \tag{2.30}$$

where, again, $\times$ denotes element-by-element multiplication.

In the benchmark case featuring early resolution of uncertainty, the implied physical probabilities are close to their estimates in the data, and the SDF correctly identifies bad states as the ones with the highest risk compensation. This is consistent with the intuition from our economic model, as shown in (2.20). Notably, there is less variation in the SDF across states, which can be explained by a lower persistence of the aggregate states in the data relative to the economic model. In the case of expected utility, implied physical probabilities are quite different from their estimates in the data. The recovered probability of bad events are so large that the implied risk compensation in bad states is actually considerably smaller than in good states: the magnitude of the SDF going to the bad state is more than two times smaller than going to the good state, while the opposite is the case for recursive utility. To obtain the economically plausible implication that the bad state requires higher risk compensation in the data than the good state, the utility structure should incorporate a sufficient degree of preference for early resolution of uncertainty. In our case, this requires $\theta$ to be below -8.

The measurements of the physical probabilities have direct implications for the moments of stock returns and macroeconomic variables. As we show in the bottom left panel of Figure 2.6, using the implied physical probabilities under expected utility leads to negative estimates of average returns of about -15%. This is a direct consequence of assigning a large probability to bad states with low negative returns. Using negative values of $\theta = -11.28$ results in a more plausible estimate of average returns of about 3%, which is more consistent with the evidence in the data. We further evaluate the implications for the measurements of physical probabilities on higher order moments of returns. The measurements of physical probabilities do not have a significant effect on the implied physical volatility of returns: it is 12.92% under expected utility, relative to 12.61% under early resolution of uncertainty with $\theta$ of -11.28. However, both skewness and kurtosis of returns are closer to the actual data under the recursive utility structure: for example, the return skewness is about -0.7 for $\theta = -11.28$ relative to 0.6 under the expected utility.

In sum, consistent with our economic model, our empirical findings suggest that the recovery of physical probabilities and the SDF is significantly affected by the underlying preference for the timing of the resolution of uncertainty. When the magnitude of the preference for early resolution of uncertainty is not fully accounted for, the implied physical probabilities tend to trade-off higher probabilities of bad economic events for lower risk compensation in these events. A specification with enough preference for early resolution of uncertainty matches the actual physical probabilities and the moments of returns in the data quite well.

## 2.5. Robustness

### *2.5.1. Measurements of Wealth-to-Consumption Ratio*

In our benchmark case, we estimate the log wealth-to-consumption ratio assuming it is proportional to the log price-dividend ratio, $\log PC \approx \alpha + \beta \log PD$, and the coefficient $\beta$ is identified from the volatility of the wealth-to-consumption ratio in the model relative to the volatility of price-dividend ratio in the data. To check the robustness of our findings,

we consider alternative values of $\beta$, and plot the implied physical probability of bad states for a range of $\beta$ and $\theta$ values on Figure 2.7. As a shown in the plot, to match a calibrated probability of the bad state of 0.25 for values of $\beta > 0$, which is the economically plausible case, the implied preference parameter $\theta$ should be below one. While the actual value of the preference parameter depends on the choice of $\beta$, for all configurations of $\beta > 0$ the implied preference structure suggests a preference for early resolution of uncertainty.

*2.5.2. Measurements of Aggregate States*

Our benchmark specification features three economic states which are identified using the 25th and 50th percentile cut-offs for the return distribution in the data. We consider various robustness checks with respect to the location and the number of bins, and we report the results for the recovery of physical probabilities and the SDF under alternative identification of the aggregate states in Table 2.4.

Specifically, we first entertain the case where the cut-off point for the bad state corresponds to the 20th percentile of the return distribution. As shown in Table 2.4, the recovery under expected utility still over-estimates the probability of the bad state to be 40% which results in a -8.47% estimate for the average market return. The value of $\theta = -11.28$ which corresponds to the preference for early resolution of uncertainty minimizes the KL divergence criterion and leads to positive average market returns and the bad state probability closer to the data. Similarly, when the left tail is set above our benchmark specification to the 30th percentile of the distribution of market returns, the implied probability of the bad state is 50% and average market returns are negative under the expected utility. On the other hand, the implied probability of the bad state is 32% while average market returns are 5.7% under the recursive utility structure with preference for early resolution of uncertainty. Note that in both of these configurations we can still meaningfully identify the aggregate economic states. For example, for the 20th percentile left tail specification, the PD ratio in the best state is 55 relative to 52 in the worst state, consumption growth in the best state is 1.8% versus 0.3% in the worst state, and the average VIX is 20 in the

best state compared to 35 in the worst state. Similar results hold for the 30th percentile for the specification of the left tail. Interestingly, as the bad state corresponds more and more to the tail events (20th percentile compared to 25th and 30th percentile), the implied value of $\theta$ becomes more negative. This suggests that the recursive utility structure and the preference for early resolution of uncertainty play an increasingly important role to account for the larger tail events in stock markets; see e.g. **?**, Eraker and Shaliastovich (2008) and Shaliastovich (2010) for the discussion of equilibrium recursive utility models with jumps.

While our benchmark specification is a three state model, similar results hold for a two state model as well, which we report in Table 2.4. We consider as a robustness check a two-state specification where the left tail (bad state) is defined as, respectively, the 20th, 25th, and 30th percentiles of quarterly returns. For all cases, an expected utility specification over-estimates the probability of bad states, and in all the cases except the 30% cut-off, the implied average market returns are negative. The value of $\theta$ that minimizes KL divergence between the recovered conditional distribution and the transition probabilities in the data are all negative and imply a preference for early resolution of uncertainty. Under recursive utility, both the probability of bad states and average market returns are much closer to the data compared to the case of expected utility.

We have also considered the robustness of our results with respect to the alternative specifications for the good state. Generally, if the bin structure permits us to meaningfully identify aggregate economic states, our results remain robust.[42]

*2.5.3. Monthly Horizon*

While our main results are presented on a quarterly frequency, our results are robust to using a monthly horizon as well.

We use options with one month to maturity and track their prices on the same expiration

[42]For a sufficiently high cut-off point used to define a good state, the relative magnitude of economic variables in good states versus bad states are reversed compared to the benchmark case. This might be due to microstructure and data issues associated with the right tail of the return distribution.

cycle as in our benchmark specification, that is, we use quarterly observations in March, June, September, and December. We construct the risk-neutral distribution implied by the monthly options prices and compute Arrow-Debreu prices following our discussion in Section 1. Economic states are defined using the capital gains to the index over the past month and are binned at the 25th and 50th percentiles, as in our benchmark setup.

We report the results for the recovery of physical probabilities and the SDF for the monthly horizons in Table 2.4, and show the implied unconditional probability of being in the bad state, the implied value of the SDF for transitions from the good to the bad state, the implied average market returns and the Kullback-Leibler divergence criterion for each value of $\theta$ in Figure 2.8. As seen in the Figure and the Table, the Kullback-Leibler divergence is minimized at $\theta = -39.46$. Under expected utility, the probability of the bad state is significantly biased upwards and is equal to 62%, and the implied average market return is very negative. Under a preference for early resolution of uncertainty, the probability of the bad state is 31% and the average market return is 6%. Overall, our evidence at the monthly horizon is similar to the benchmark specification and the economic model.

## 2.6. Conclusion

We show how to separately recover physical probabilities and risk adjustments from the risk-neutral probabilities without using macroeconomic variables and allowing for a preference for timing of resolution of uncertainty thus extending the recovery framework of Ross (2011) to the Kreps and Porteus (1978) recursive preferences of Epstein and Zin (1989) and Weil (1989). We implement our market-based recovery framework using S&P500 options and find that the data strongly support a specification of early resolution of uncertainty, with preference parameter values similar to common values in the literature. Using the data and model simulations, we document significant biases in estimating physical probabilities and risk adjustments when the preference for early resolution of uncertainty is not sufficiently accounted for.

To highlight the implications of timing of resolution of uncertainty for the physical probabilities and risk adjustments, we first use in a Mehra and Prescott (1985)-Weil (1989) economic model, which incorporates Epstein-Zin utility. We calibrate our model to match stylized facts of financial market returns such as the equity premium and average risk-free rates. In the model, we see that failing to sufficiently account for a preference for early resolution of uncertainty leads to biased estimates of the physical distribution of returns, because we will attribute too large a proportion of the high state prices of bad states to physical likelihoods rather than risk-adjustment.

We then implement our market-based recovery approach using S&P500 options data. We extract the risk-neutral distribution from options prices using a standard technique, and identify our economic states based on market returns. We apply the framework to the extracted risk-neutral distribution and recover implied dynamics for the physical return probabilities of the U.S. market. We show that not fully accounting for the preference for early resolution of uncertainty results in an over-estimation of the probabilities of bad states and downward-biased estimates of average returns. In all, the evidence from the S&P500 index options market suggests that the representative agent for the U.S. economy has a strong preference for early resolution of uncertainty.

Tables

Table 2.1: Model Calibration

| Preferences | $\delta$ | $\gamma$ | $\psi$ |
|---|---|---|---|
| | $0.988^*$ | 25 | 2 |
| Consumption | $\mu$ | $\rho$ | $\sigma$ |
| | $1.88^*$ | $0.28^*$ | $1.32^*$ |
| Dividend | $\mu_d$ | $\phi$ | |
| | $1.88^*$ | 3 | |

Calibration of model parameters. The model is calibrated on a quarterly frequency. The parameter values with superscript $^*$ are annualized, e.g. $\delta^4$ and $\rho^4$ for the subjective discount factor and consumption growth persistence, $2\sigma$ for consumption volatility, and $4\mu$ for the mean. The AR(1) dynamics of consumption growth is discretized into 3 states using the Tauchen and Hussey (1991) quadrature approach. Mean and volatility parameters are in percent.

Table 2.2: Model Output

| | Data | Model Under EZ | Model Under EU |
|---|---|---|---|
| $\mathbb{E}[\frac{C_{t+1}}{C_t}]$ | 1.90 | 1.90 | -0.77 |
| $\sigma\left[\frac{C_{t+1}}{C_t}\right]$ | 2.20 | 2.50 | 1.90 |
| $\rho\left[\frac{C_{t+1}}{C_t}\right]$ | 0.50 | 0.50 | 0.44 |
| $\mathbb{E}[PD_m]$ | 60.02 | 59.60 | 58.16 |
| $\mathbb{E}[R_m]$ | 7.13 | 8.89 | 0.90 |
| $\mathbb{E}[Rf]$ | 1.19 | 1.54 | 0.87 |

Data and model-implied mean, volatility, and persistence of annual consumption growth (top panel), and average price-dividend ratio, excess returns on the market and the risk-free rate (bottom panel). Data is annual from 1929 to 2010; model statistics are based on a long simulation of quarterly data time-aggregated to an annual horizon. "Under EZ" model output is based on the recursive utility configuration with preference for early resolution of uncertainty, while "Under EU" model output is based on the implied physical probabilities recovered under the expected utility assumption.

Table 2.3: Economic Variables in Aggregate States

| | Mean | Std. Dev. | Econ. State Good | Medium | Bad |
|---|---|---|---|---|---|
| Mkt Capital Gains | 5.92 | 17.82 | 22.63 | 0.19 | -35.51 |
| Mkt PD ratio | 57.81 | 14.00 | 54.50 | 57.29 | 49.54 |
| VIX | 22.31 | 8.08 | 20.22 | 18.77 | 30.29 |
| Real cons. growth | 1.31 | 0.83 | 1.79 | 1.74 | 0.29 |
| RN Distribution: | | | | | |
| Volatility | | | 18.92 | 19.57 | 25.27 |
| 3rd Moment×1000 | | | -0.74 | -0.84 | -1.24 |
| 4rd Moment×1000 | | | 0.37 | 0.39 | 0.73 |

The top panel shows the mean and standard deviation of asset-price and macroeconomic variables, and their median values in good, medium, and bad economic states. Bottom panel shows the volatility, and 3rd and 4th centered moments of the risk-neutral distribution in each state. Economic states correspond to upper 50%, 25%-50%, and lower 25% distribution of capital gains of S&P500, respectively. Quarterly observations from 1996 to 2011.

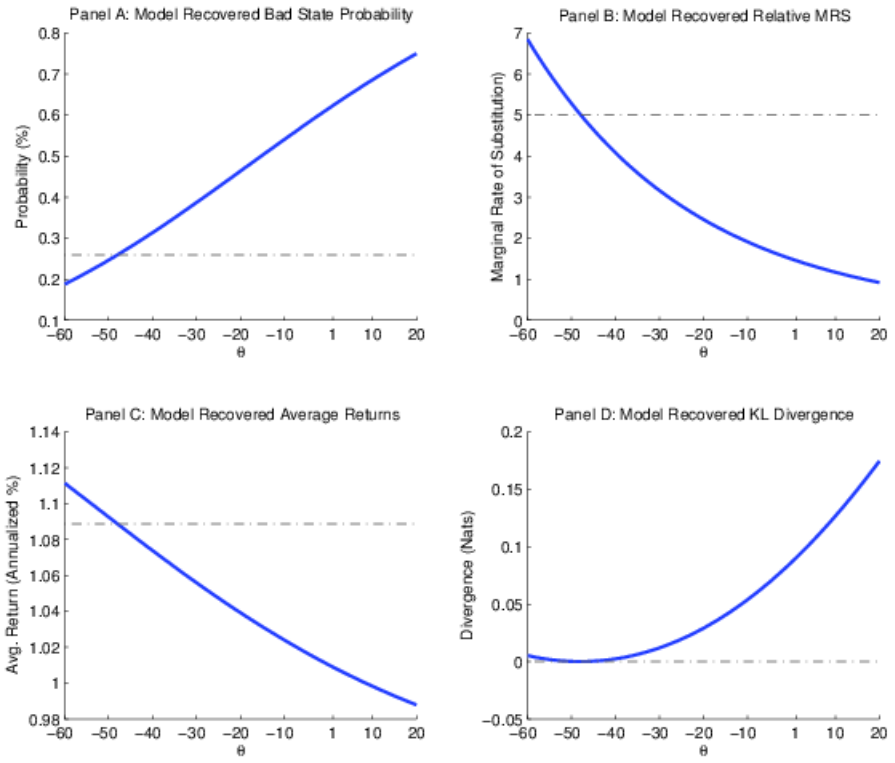
Table 2.4: Implications for Probabilities and Risk Compensations

| | Under EZ | | | | Under EU | | | |
|---|---|---|---|---|---|---|---|---|
| | $\theta$ | Pr. Bad | SDF | Mkt Ret | $\theta$ | Pr. Bad | SDF | Mkt Ret |
| Benchmark | -11.28 | 0.28 | 1.57 | 2.56 | 1.00 | 0.59 | 0.37 | -14.85 |
| Three States: | | | | | | | | |
| *Left Tail:* | | | | | | | | |
| 20$th$ pctile | -12.21 | 0.25 | 1.23 | 1.33 | 1.00 | 0.39 | 0.54 | -8.47 |
| 30$th$ pctile | -7.27 | 0.32 | 1.32 | 5.90 | 1.00 | 0.48 | 0.63 | -0.62 |
| Two States: | | | | | | | | |
| *Left Tail:* | | | | | | | | |
| 20$th$ pctile | -52.34 | 0.21 | 1.52 | 1.52 | 1.00 | 0.27 | 1.07 | -2.15 |
| 25$th$ pctile | -9.43 | 0.25 | 1.87 | 1.94 | 1.00 | 0.35 | 1.19 | -2.88 |
| 30$th$ pctile | -3.09 | 0.30 | 1.47 | 5.65 | 1.00 | 0.34 | 1.25 | 4.35 |
| Monthly horizon | -39.46 | 0.31 | 1.12 | 6.07 | 1.00 | 0.62 | 0.22 | -26.73 |

Implied physical probability of a bad aggregate state, the value of the stochastic discount factor from good to bad state relative to staying in a good state, and the implied average annualized market return, recovered under the specifications with recursive preferences ("Under EZ") and under power utility ("Under EU"). Recursive preference model output corresponds to the optimal preference parameter $\theta$ which minimizes Kullback-Leibler divergence between the implied and physical conditional probabilities of aggregate states in the data; under expected utility, $\theta$ is fixed at 1. The benchmark setup features 3 states and the bad state cut-off at 25% of the return distribution. Robustness specifications include setting the bad state cut-off to 20th and 30th percentiles; using two state, and using monthly data horizons. Quarterly observations from 1996 to 2011.

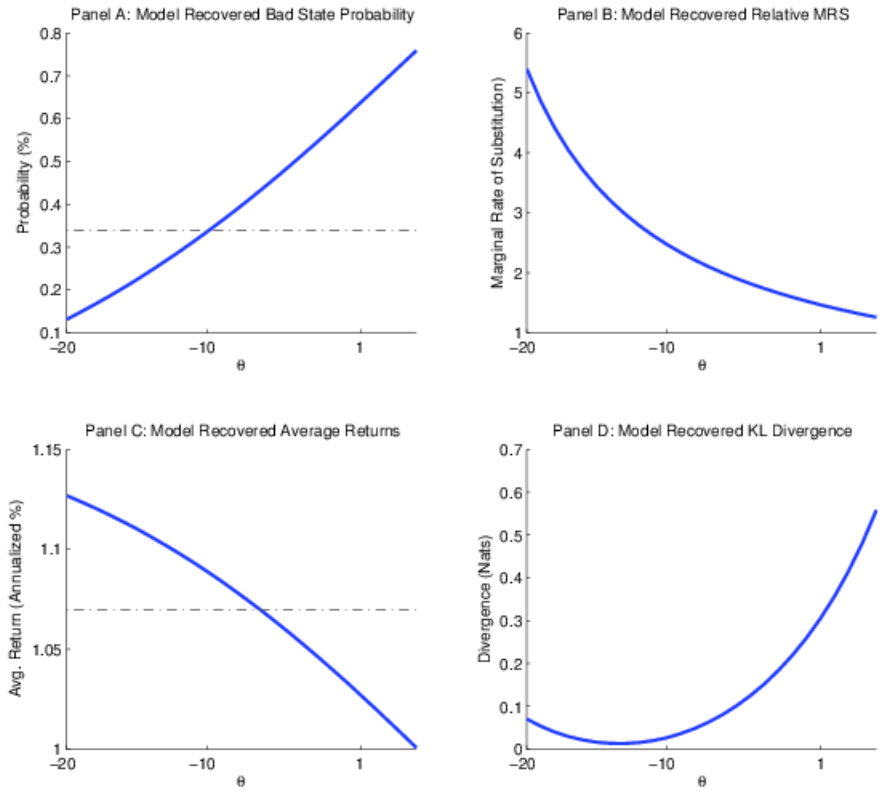
## Figures

Figure 2.1: Probabilities and Risk Adjustments: Model



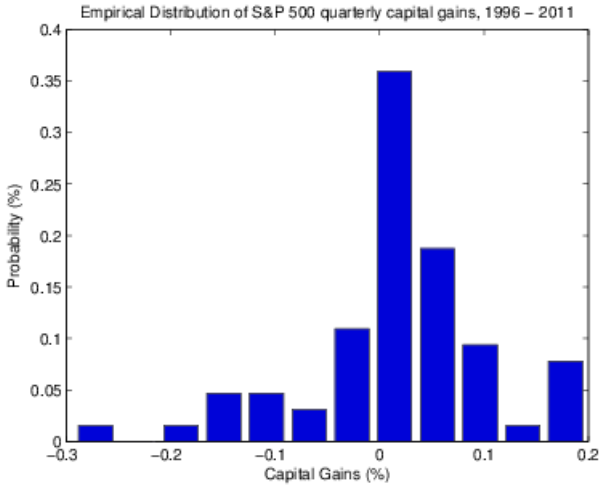

Implications for physical probabilities and risk adjustments in the economic model using consumption growth states. Top panel shows unconditional probability of bad states recovered in the model and the value of the stochastic discount factor from good to bad state relative to staying in a good state as a function of the preference parameter $\theta$. Bottom panel shows the implied average market return and the Kullback-Leibler divergence between the implied and calibrated physical probabilities as a function of the preference parameter $\theta$. The dashed line represents the value given by $\theta$ that minimizes the Kullback-Leibler (KL) divergence. The output is based on the economic model.

Figure 2.2: Probabilities and Risk Adjustments: Model, Return States



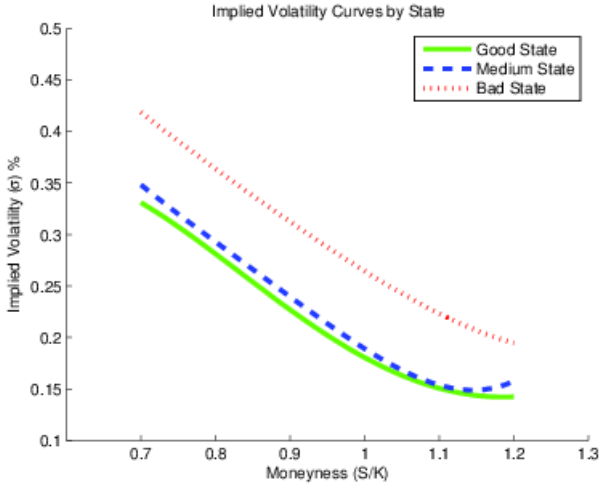

Implications for physical probabilities and risk adjustments in the economic model using return states. Top panel shows unconditional probability of bad states recovered in the model and the value of the stochastic discount factor from good to bad state relative to staying in a good state as a function of the preference parameter $\theta$. Bottom panel shows the implied average market return and the Kullback-Leibler divergence between the implied and calibrated physical probabilities as a function of the preference parameter $\theta$. The dashed line represents the value given by $\theta$ that minimizes the Kullback-Leibler (KL) divergence. The output is based on the economic model.

Figure 2.3: Empirical Distribution of Market Capital Gains



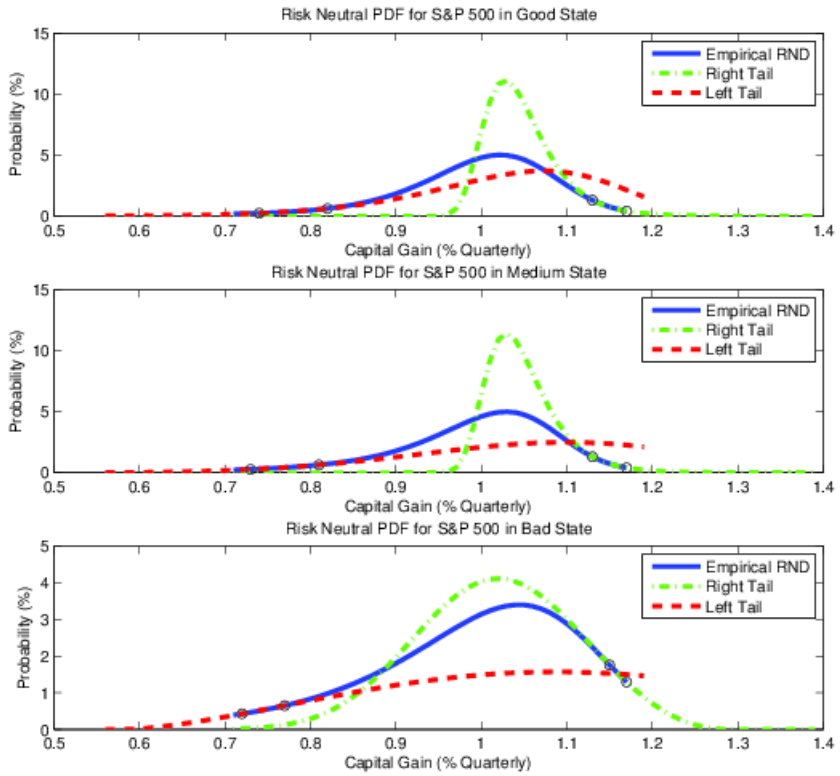

Empirical distribution of log capital gains on S&P500. Quarterly observations from 1996 to 2011.

Figure 2.4: Implied Volatility Curves in Economic States



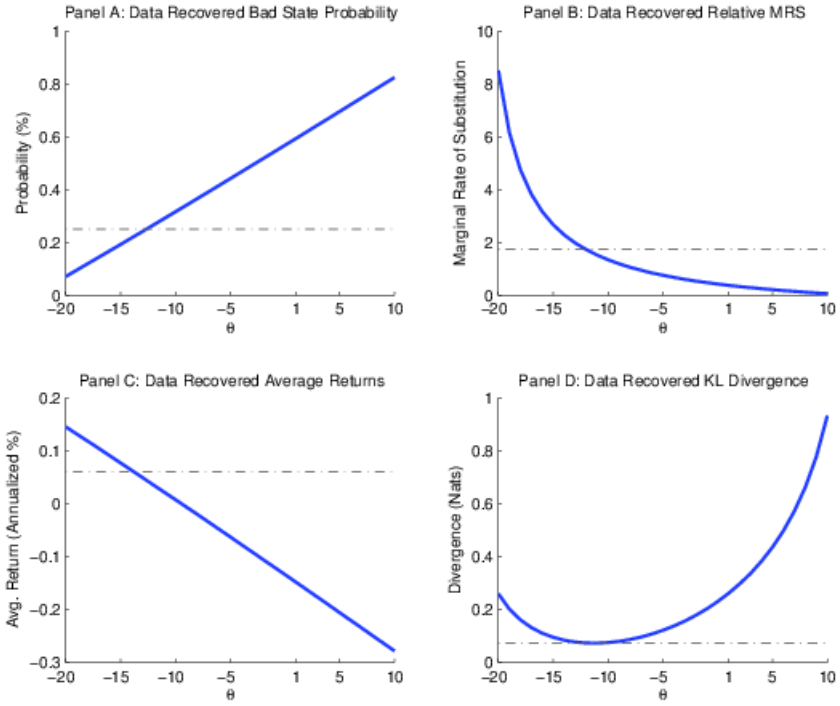

Implied volatility curves for a range of moneyness (spot/strike) in each aggregate economic state. Quarterly observations from 1996 to 2011.

Figure 2.5: Recovered Risk-Neutral Densities in Economic States



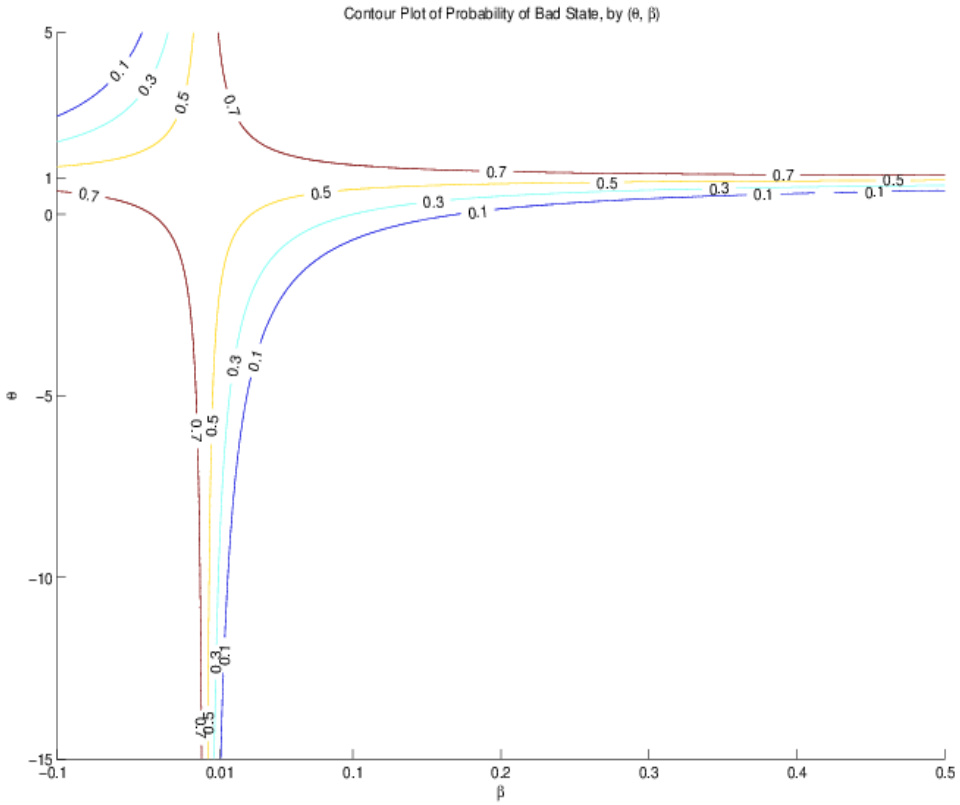

Empirical risk-neutral densities for the market capital gains in good, medium and bad economic states. The blue solid line represents the portion constructed from option data alone; red dashed and greed dashed-dotted lines represent left and right tails, respectively, constructed using the GEV approximation to the underlying data density. Quarterly observations from 1996 to 2011.

Figure 2.6: Probabilities and Risk Adjustments: Data



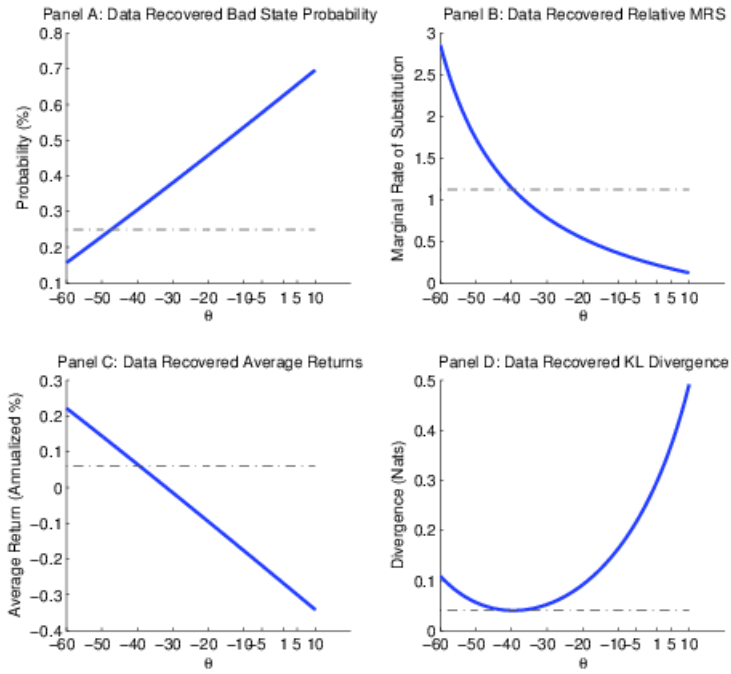

Implications for physical probabilities and risk adjustments in the data. Top panel shows unconditional probability of bad states recovered in the model and the value of the stochastic discount factor from good to bad state relative to staying in a good state as a function of the preference parameter $\theta$. Bottom panel shows the implied average market return and the Kullback-Leibler divergence between the implied and calibrated physical probabilities as a function of the preference parameter $\theta$. The dashed line represents the value given by $\theta$ that minimizes the Kullback-Leibler (KL) divergence. The output is based on quarterly observations from 1996 to 2011.

Figure 2.7: Implied Bad State Probability




Contour plot of recovered bad state probability in the data, for different combinations of preference parameters $\theta$ (y-axis) and wealth-to-consumption scale factor $\beta$ (x-axis). Quarterly observations from 1996 to 2011.

Figure 2.8: Implied Probabilities and Risk Adjustments: Monthly Data




Top panel shows unconditional probability of bad states recovered in the model and the value of the stochastic discount factor from good to bad state relative to staying in a good state as a function of the preference parameter $\theta$. Bottom panel shows the implied average market return and the Kullback-Leibler divergence between the implied and calibrated physical probabilities as a function of the preference parameter $\theta$. The dashed line represents the value given by $\theta$ that minimizes the Kullback-Leibler (KL) divergence. The output is based on monthly horizons at primary cycle expirations (March - June - September - December) from 1996 to 2011.

# CHAPTER 3 : Volatility-of-Volatility Risk

(with Ivan Shaliastovich)

## 3.1. Introduction

Recent studies show that volatility risks significantly affect asset prices and the macroeconomy.[43] In the data, asset market volatility can be directly captured by the volatility index (VIX). Calculated from the cross-section of S&P500 option prices, the VIX index provides a risk-neutral forecast of the aggregate index volatility over the next 30 days. The VIX index exhibits substantial fluctuations, which in the data and in many economic models drive the movements in asset prices and risk premia. Interestingly, the volatility of the VIX index itself varies over time. Computed from VIX options in an analogous way to the VIX, the volatility-of-volatility index (VVIX) directly measures the risk-neutral expectations of the volatility of volatility in the financial markets. In the data, we find that the VVIX has separate dynamics from the VIX, so that fluctuations in volatility of volatility are not directly tied to movements in market volatility. The volatility-of-volatility risks are a significant risk factor which affects the time-series and the cross-section of index and VIX option returns, above and beyond volatility risks. The evidence in the data is consistent with a no-arbitrage model which features time-varying market volatility and volatility-of-volatility factors which are priced by the investors. In particular, volatility and volatility of volatility have negative market prices of risk, so that investors dislike increases in volatility and volatility of volatility, and demand a risk compensation for the exposure to these risks.

Our no-arbitrage model follows and extends the one-factor stochastic volatility specification of equity returns in Bakshi and Kapadia (2003). Specifically, we introduce a separate time-varying volatility-of-volatility risk factor which drives the conditional variance of the variance of market returns.[44] Both factors are priced in our model. We use the model to

[43] See e.g. Bansal and Yaron (2004), Bloom (2009), Bansal et al. (2013), Fernandez-Villaverde and Rubio-Ramírez (2013) for the discussion of macroeconomic volatility risks, and Coval and Shumway (2001), Bakshi and Kapadia (2003), Campbell, Giglio, Polk, and Turley (2012) for market volatility risks.

[44] We use the terms "variance risk" and "volatility risk" interchangeably unless otherwise specified.

characterize the payoffs to delta-hedged equity and volatility options. The zero-cost, delta-hedged positions represent the gains on a long position in the option, continuously hedged by an offsetting short position in the underlying asset. As argued in Bakshi and Kapadia (2003), delta-hedged option payoffs are very useful to study volatility-related risks as they most cleanly isolate the exposures to volatility risks.[45] Indeed, under a standard linear risk premium assumption, we show that the expected payoff on the delta-hedged position in equity index options consists of the risk compensations for both volatility and volatility-of-volatility risks. For volatility options, the expected payoffs only involve the compensation for volatility-of-volatility risks. The risk compensations are given by the product of the market price of risk, the risk exposure of the asset, and the time-varying quantity of each source of risk. The model thus delivers clear, testable predictions for the expected option returns and their relation to volatility and volatility-of-volatility risks. In the model, if investors dislike volatility and volatility of volatility so that the market prices of these risks are negative, delta-hedged equity and VIX option gains are negative on average. In the cross-section, the average returns are more negative for option strategies which have higher exposure to the volatility and volatility-of-volatility risks. Finally, in the time series, higher volatility and volatility of volatility predict more negative delta-hedged option gains in the future.

To evaluate model implications for volatility and volatility risks, we use monthly observations on the implied and realized variances for the market index and the VIX, and index and VIX option price data over the 2006 - 2013 sample. We verify that the option-implied volatility measures capture meaningful economic information about the uncertainty in future market returns and market volatility in the data. Using predictive regressions, we show that the VIX is a significant predictor of the future realized variance of market returns, while the VVIX significantly forecasts future realized variation in the VIX index itself. Including both volatility measures at the same time, we find that the predictive power is

[45]For example, unlike delta-hedged positions, zero-beta straddles analyzed in Coval and Shumway (2001) are not dynamically rebalanced and may contain a significant time-decay option premium component.

concentrated with the corresponding factor (i.e., the VIX for market return volatility, the VVIX for VIX volatility), and the other variable has an insignificant impact. This evidence confirms that the measured VIX and VVIX indices can indeed separately capture volatility and volatility-of-volatility movements in the asset markets.

In the time-series, the VVIX behaves quite differently from the VIX, consistent with a setup of our model which separates market volatility from volatility of volatility. The VVIX is much more volatile, and is less persistent than the VIX. The correlation between the two series is 0.30. While both volatility measures share several common peaks, most notably during the financial crisis, other times of economic distress and economic uncertainty, such as the Eurozone debt crisis and flash crash in May 2010 and the U.S. debt ceiling crisis in August 2011, are characterized by large increases in the VVIX with relatively little action in the VIX. On average, the risk-neutral volatilities of the market return and market volatility captured by the VIX and VVIX exceed the realized volatilities of returns and the VIX. The difference between the risk-neutral and physical volatilities of market returns is known as the variance premium (variance-of-variance premium for the VIX), and the findings of positive variance and variance-of-variance premium suggest that investors dislike variance and variance-of-variance risks, and demand a premium for being exposed to these risks.

We next turn to the asset-price evidence from the equity index and VIX option markets. In line with our model, we consider discrete-time counterparts to the continuously-rebalanced delta-hedged gains; this approach is similar to Bakshi and Kapadia (2003). Consistent with the evidence in previous studies, the average delta-hedged returns on out-of-the-money equity index calls and puts are significantly negative in our sample. The novel evidence in our paper is that the average delta-hedged returns on VIX options are also negative and statistically significant at all strikes, except for out-of-the-money puts which are marginally significant. Estimates of the loss for call options range from -0.57% of the index value for in-the-money VIX calls to -1.41% for out-of-the-money calls. The negative average returns on index and VIX options directly suggest that the market prices of volatility and

volatility-of-volatility risks are negative.

We then show that the cross-sectional spreads in average option returns are significantly related to the volatility and volatility-of-volatility risks. In lieu of calculating exact model betas, we compute proxies for the option exposures to the underlying risks using the Black and Scholes (1973b) vega and volga. Vega represents an increase in the Black-Scholes value of the option as the implied volatility increases by 1%, and thus provides an estimate for the exposure of equity options to volatility risks, and of VIX options to volatility-of-volatility risks. Volga is the second partial derivative of the option price with respect to the volatility, which we use to measure the sensitivity of the index option price to the volatility-of-volatility risks. Vega and volga vary intuitively with the moneyness of the option in the cross-section, and help us proxy for the betas of the options to the underlying risks. Empirically, we document that average option returns are significantly and negatively related to our proxies for volatility and volatility-of-volatility risks. Hence, using the cross-section of equity index options and VIX options, we find strong evidence for a negative market price of volatility and volatility-of-volatility risks.

Finally, we consider a predictive role of our volatility measures for the future option returns. In the model, expected delta-hedged gains are time-varying and are driven by the volatility and volatility of volatility (by volatility-of-volatility for VIX options). In particular, as option betas are all positive, when the market prices of volatility-related risks are negative, both volatility measures should forecast future returns with a negative sign. This model prediction is consistent with the data. The VIX and VVIX significantly negatively predict future index option returns, and the VVIX is a significant negative predictor of option returns on the VIX. Hence, using the cross-sectional and time-series evidence from the option markets, we find strong support that both volatility and volatility-of-volatility risks are separate priced sources of in the option markets, and have negative market prices of risks.

**Related Literature.** Our paper is most closely related to Bakshi and Kapadia (2003) who

consider the implications of volatility risk for equity index option markets. We extend their approach to include volatility-of-volatility risk, and bring evidence from VIX options. To help us focus on the volatility-related risks, we consider dynamic delta-hedging strategies where a long position in option is dynamically hedged by taking an offsetting position in the underlying. Delta-hedged strategies are also used in Bertsimas, Kogan, and Lo (2000), Cao and Han (2013) and Frazzini and Pedersen (2012), and are a standard risk management technique of option traders in the financial industry (Hull (2011)). In an earlier study, Coval and Shumway (2001) considers the returns on zero-beta straddles to identify volatility risk sensitive assets. Zhang and Zhu (2006) and Lu and Zhu (2010) highlight the nature and importance of volatility risks by analyzing the pricing of VIX futures. Also notably our analysis suggests that variance dynamics are richer than that of the square-root process typically considered in the literature - these findings are consistent with the results of Christoffersen, Jacobs, and Mimouni (2010) and Branger, Kraftschik, and Völkert (2014).

In a structural approach, Bollerslev et al. (2009) consider a version of the Bansal and Yaron (2004) long-run risks model which features recursive utility and fluctuations in the volatility and volatility of volatility of the aggregate consumption process. They show that in equilibrium, investors require compensation for the exposure to volatility and volatility-of-volatility risks. With preference for early resolution of uncertainty, the market prices of the two risks are negative. As a result, the variance risk premium is positive on average, and can predict future equity returns. Bollerslev et al. (2009) and Drechsler and Yaron (2011) show that the calibrated version of such a model can account for the key features of equity markets and the variance premium in the data. Our empirical results in the paper are consistent with the economic intuition in these models and complement the empirical evidence in these studies.

Finally, it is worth noting that in our paper we abstract from jumps in equity returns, and focus on diffusive volatilities as the main drivers of asset prices and risk premia. For robustness, we confirm that our predictability results are robust to controlling for jump

risk measures such as the slope of the implied volatility curve, realized jump intensity (Barndorff-Nielsen and Shephard (2006) and Wright and Zhou (2009)), and risk-neutral skewness (Bakshi et al. (2003)). Hence, we argue that the VIX and VVIX have a significant impact on option returns even in the presence of stock market and volatility jumps; we leave a formal treatment of jumps for future research. Reduced-form models which highlight the role of jumps include Bates (2000), Pan (2002), and Duffie et al. (2000), among others.

Our paper proceeds as follows. In Section 2 we discuss our model which links expected delta-hedged equity and volatility option gains to risk compensations for volatility and volatility-of-volatility risk. In Section 3, we describe the construction of both the model-free implied variance measures and high-frequency realized variance measures, and summarize their dynamics in the time-series. We show that the implied variances have a strong ability to forecast future realized variance. Section 4 provides the empirical evidence from option prices by empirically implementing the delta-hedged option strategies in our model. Section 5 presents robustness tests for alternative measures of variance, as well as robustness of the results in the presence of jump risks. Section 6 concludes the paper.

## 3.2. Model

In this section we describe our model for stock returns, and both equity and volatility option prices. Our model is an extension of Bakshi and Kapadia (2003) and features time-varying market volatility and volatility-of-volatility factors. Both volatility risks are priced, and affect the level and time-variation of the expected asset payoffs.

### *3.2.1. Dynamics of Equity and Equity Option Prices*

Under the physical measure ($\mathbb{P}$), the stock price $S_t$ evolves according to:

$$\begin{aligned} \frac{dS_t}{S_t} &= \mu(S_t, V_t, \eta_t)dt + \sqrt{V_t}dW_t^1, \\ dV_t &= \theta(V_t)dt + \sqrt{\eta_t}dW_t^2, \\ d\eta_t &= \gamma(\eta_t)dt + \phi\sqrt{\eta_t}dW_t^3, \end{aligned} \tag{3.1}$$

where $dW_t^i$ are the Brownian motions which drive stock returns, the stock return variance, and the variance of the variance, for $i = 1, 2, 3$, respectively. The Brownian components can be correlated: $dW_t^i \, dW_t^j = \rho_{i,j} \, dt$ for all $i \neq j$. $V_t$ is the variance of instantaneous returns and $\eta_t$ is the variance of innovations in $V_t$. Note that the drift of the variance $V_t$ only depends on itself, and not not on the returns $S_t$ or the volatility of volatility $\eta_t$. Similarly, the drift of the volatility of volatility $\eta_t$ is a function of the volatility of volatility $\eta_t$.

Under the risk-neutral measure ($\mathbb{Q}$), the stock price $S_t$ follows a similar process, where the drifts are adjusted by the risk compensations for the corresponding risks:

$$\begin{aligned} \frac{dS_t}{S_t} &= r_f dt + \sqrt{V_t} d\tilde{W}_t^1, \\ dV_t &= (\theta(V_t) - \lambda_t^V)dt + \sqrt{\eta_t} d\tilde{W}_t^2, \\ d\eta_t &= (\gamma(\eta_t) - \lambda_t^\eta)dt + \phi\sqrt{\eta_t} d\tilde{W}_t^3. \end{aligned} \tag{3.2}$$

In this representation, the $\tilde{W}_t^i$ represent Brownian motions under the risk-neutral $\mathbb{Q}$ measure. $\lambda_t^V$ captures the risk compensation for the variance risk, and $\lambda_t^\eta$ reflects the compensation for the innovations in the the variance of variance. If investors dislike variance and variance-of-variance risks, the two risk compensations are negative. In this case, the variances have higher drifts under the risk-neutral measure than under the physical measure.

Let $C_t(K, \tau)$ denote the time $t$ price of a call option on the stock with strike price $K$ and time to maturity $\tau$. Assume the risk-free rate $r_f$ is constant. To simplify the presentation, we further abstract from dividends. While we focus our discussion on call options, the case of put options follows analogously. Given the specified dynamics of the stock price under the two probability measures, the option price is given by a twice-differentiable function $C$ of the state variables: $C_t(K, \tau) = C(S_t, V_t, \eta_t, t)$. By Itô's Lemma,

$$dC_t = \frac{\partial C}{\partial S} dS_t + \frac{\partial C}{\partial V} dV_t + \frac{\partial C}{\partial \eta} d\eta_t + b_t dt, \tag{3.3}$$

for a certain drift component $b_t$.

The discounted option price $e^{-r_f t}C_t$ is a martingale under $\mathbb{Q}$ and thus has zero drift. We use Itô's Lemma again to obtain that:

$$\frac{\partial C}{\partial S}S_t r_f + \frac{\partial C}{\partial V}\left(\theta(V_t) - \lambda_t^V\right) + \frac{\partial C}{\partial \eta}\left(\gamma(\eta_t) - \lambda_t^\eta\right) + b_t - r_f C_t = 0. \tag{3.4}$$

This implies that:

$$b_t = r_f\left(C_t - \frac{\partial C}{\partial S}S_t\right) - \frac{\partial C}{\partial V}(\theta(V_t) - \lambda_t^V) - \frac{\partial C}{\partial \eta}(\gamma(\eta_t) - \lambda_t^\eta). \tag{3.5}$$

Let $\Pi_{t,t+\tau}$ stand for the delta-hedged option gain for call options:

$$\Pi_{t,t+\tau} \equiv C_{t+\tau} - C_t - \int_t^{t+\tau}\frac{\partial C}{\partial S}dS_u - \int_t^{t+\tau} r_f\left(C_u - \frac{\partial C}{\partial S}S_u\right)du. \tag{3.6}$$

The delta-hedged option gain represents the gain on a long position in the option, continuously hedged by an offsetting short position in the stock, with the net balance earning the risk-free rate.

Combining equations (3.5) and (3.3) together, we obtain that the delta-hedged option gain for call options is given by,

$$\begin{aligned}\Pi_{t,t+\tau} &\equiv C_{t+\tau} - C_t - \int_t^{t+\tau}\frac{\partial C}{\partial S}dS_u - \int_t^{t+\tau} r_f\left(C_u - \frac{\partial C}{\partial S}S_u\right)du \\ &= \int_t^{t+\tau}\lambda_u^V\frac{\partial C}{\partial V}du + \int_t^{t+\tau}\lambda_u^\eta\frac{\partial C}{\partial \eta}du + \int_t^{t+\tau}\frac{\partial C}{\partial V}\sqrt{\eta_u}dW_u^2 + \int_t^{t+\tau}\frac{\partial C}{\partial \eta}\phi\sqrt{\eta_u}dW_u^3.\end{aligned} \tag{3.7}$$

Since the expectation of Itô integrals is zero, the expected delta-hedged equity option gains are given by:

$$\mathbb{E}_t\left[\Pi_{t,t+\tau}\right] = \mathbb{E}_t\left[\int_t^{t+\tau}\lambda_u^V\frac{\partial C}{\partial V}du\right] + \mathbb{E}_t\left[\int_t^{t+\tau}\lambda_u^\eta\frac{\partial C}{\partial \eta}du\right]. \tag{3.8}$$

The expected option gains depend on the risk compensation components for the volatility and volatility-of-volatility risks ($\lambda_t^V$ and $\lambda_t^\eta$), and the option price exposures to these two sources of risks ($\frac{\partial C}{\partial V}$ and $\frac{\partial C}{\partial \eta}$). For tractability and consistency with the literature, we assume that the risk premium structure is linear:

$$\lambda_t^V = \lambda^V V_t, \quad \lambda_t^\eta = \lambda^\eta \eta_t, \tag{3.9}$$

where $\lambda^V$ is the market price of the variance risk and $\lambda^\eta$ is the market price of the variance-of-variance risk. We can further operationalize (3.8) by applying Itô-Taylor expansions (Milstein (1995)). This gives us a linear factor model structure (see details in the Appendix):

$$\frac{\mathbb{E}_t\left[\Pi_{t,t+\tau}\right]}{S_t} = \lambda^V \beta_t^V V_t + \lambda^\eta \beta_t^\eta \eta_t. \tag{3.10}$$

The sensitivities to the risk factors are given by:

$$\beta_t^V = \sum_{n=0}^{\infty} \frac{\tau^{1+n}}{(1+n)!} \Phi_{t,n}^V > 0, \quad \beta_t^\eta = \sum_{n=0}^{\infty} \frac{\tau^{1+n}}{(1+n)!} \Phi_{t,n}^\eta > 0, \tag{3.11}$$

where $\Phi_{t,n}^V$ and $\Phi_{t,n}^\eta$ are positive functions which depend on the moneyness of the option and $\frac{\partial C}{\partial V}$ and $\frac{\partial C}{\partial \eta}$, respectively. Hence, the expected payoff on the delta-hedged option position combines the risk compensations for the volatility and volatility-of-volatility risks. The two risk compensations are given by the product of the market price of risk, the exposure of the asset to the corresponding risk, and the quantity of risk. In particular, options are positive-beta assets to both volatility and volatility-of-volatility risks. Hence, if investors dislike volatility and volatility-of-volatility risks so that their market prices of risks are negative, the expected option payoffs are negative as well.

*3.2.2. Dynamics of VIX Option Prices*

The squared VIX index is the annualized risk-neutral expectation of the quadratic variation of returns from time $t$ to $t+\tau$, given by:

$$VIX_t^2 = \frac{1}{\tau}\mathbb{E}_t^{\mathbb{Q}}\left[\int_t^{t+\tau} V_s \, ds\right]. \tag{3.12}$$

Given our model assumptions, the VIX index is a function of the stock market variance: $VIX_t = VIX(V_t)$. For example, in a linear model where the variance drift $\theta(V_t)$ is linear in $V_t$, the squared VIX is a linear function of the stock market variance $V_t$.

Let $F_t$ be the time $t$ price of a VIX futures contract expiring at $t+\tau$. Under no-arbitrage and continuous mark-to-market, $F_t$ is a martingale under the risk-neutral measure $\mathbb{Q}$:

$$F_t = \mathbb{E}_t^{\mathbb{Q}}\left[VIX_{t+\tau}\right] = \mathbb{E}_t^{\mathbb{Q}}\left[VIX(V_{t+\tau})\right]. \tag{3.13}$$

Under our model structure, the futures price $F$ is a function of the market variance $V_t$ and volatility of volatility $\eta$. Under economically plausible scenarios, the futures price is monotone is the two volatility processes.[46] Knowing $F_t$ and $\eta_t$ is sufficient for $V_t$, so we can re-write the economic states $\begin{bmatrix} V_t & \eta_t \end{bmatrix}$ in terms of $\begin{bmatrix} F_t & \eta_t \end{bmatrix}$.

Let $C_t^*$ be the time $t$ price of a VIX call option, whose underlying is a VIX forward contract. The option price is given by a twice differentiable function of the state variables $C^*$, so that $C_t^*(K,\tau) = C^*(F_t, \eta_t, t)$. By Itô's Lemma:

$$dC_t^* = \frac{\partial C^*}{\partial F}dF_t + \frac{\partial C^*}{\partial \eta}d\eta_t + b_t^* dt, \tag{3.14}$$

for a drift component $b_t^*$.

Under the risk-neutral measure $\mathbb{Q}$, the discounted VIX option price process $e^{-r_f t}C_t^*$ is a

[46] See also Zhang and Zhu (2006), Lu and Zhu (2010), and Branger et al. (2014) for VIX futures pricing models.

martingale, so it must have zero drift:

$$\frac{\partial C^*}{\partial F}\mathcal{D}^{\mathbb{Q}}[F_t] + \frac{\partial C^*}{\partial \eta}(\gamma(\eta_t) - \lambda_t^\eta) - r_f C_t^* + b_t^* dt = 0. \tag{3.15}$$

This implies that

$$\begin{aligned} b_t^* &= r_f C_t^* + \frac{\partial C^*}{\partial \eta}\lambda_t^\eta - \frac{\partial C^*}{\partial \eta}\gamma(\eta_t) - \frac{\partial C^*}{\partial F}\mathcal{D}^{\mathbb{Q}}[F_t] \\ &= r_f C_t^* + \frac{\partial C^*}{\partial \eta}\lambda_t^\eta - \frac{\partial C^*}{\partial \eta}\gamma(\eta_t), \end{aligned} \tag{3.16}$$

where the second line follows since $F_t$ is a martingale under $\mathbb{Q}$.

Combining the above with equation (3.14), we obtain the equation for the delta-hedged VIX option gain:

$$\begin{aligned} \Pi_{t,t+\tau}^* &= C_{t+\tau}^* - C_t^* - \int_t^{t+\tau} \frac{\partial C^*}{\partial F} dF_s - \int_t^{t+\tau} r_f C_s^* ds \\ &= \int_t^{t+\tau} \frac{\partial C^*}{\partial \eta}\lambda_s^\eta ds + \int_t^{t+\tau} \frac{\partial C^*}{\partial \eta}\phi\sqrt{\eta_s} dW_s^3. \end{aligned} \tag{3.17}$$

The delta-hedged VIX option gain in equation (3.17) is the counterpart to the delta-hedged equity option gain in equation (3.7). The difference comes from the fact that short stock position serving as the hedge in the case of equity options is funded at $r_f$ while for VIX futures the hedging position is zero cost.

Taking expectations, we can derive a corresponding expected gain on delta-hedged VIX options. Specifically, under the assumption that the risk premia are linear, we can show that (see the Appendix for details):

$$\begin{aligned} \frac{\mathbb{E}_t\left[\Pi_{t,t+\tau}^*\right]}{F_t} &= \frac{1}{F_t}\int_t^{t+\tau} \mathbb{E}_t\left[\frac{\partial C_s^*}{\partial \eta_s}\lambda_s^\eta\right] ds \\ &= \lambda^\eta \beta_t^* \eta_t. \end{aligned} \tag{3.18}$$

The delta-hedged VIX option exposure to volatility-of-volatility risks is defined as $\beta_t^* =$

$\sum_{n=0}^{\infty} \frac{\tau^{1+n}}{(1+n)!}\Phi_{t,n}^{*}$. It is a positive function, which depends on the moneyness of the option and option sensitivity $\frac{\partial C^{*}}{\partial \eta}$. Notably, while delta-hedged equity options are exposed to both the volatility $V_t$ and volatility-of-volatility risks $\eta_t$ (see equation (3.10)), delta-hedged VIX strategies are exposed only to the volatility-of-volatility risks. This helps us identify the relative importance of the two risks in the data.

## 3.3. Variance Measures

### *3.3.1. Construction of Variance Measures*

The VIX index is a model-free, forward-looking measure of implied volatility in the U.S. stock market, published by the Chicago Board Options Exchange (CBOE). The square of the VIX index is defined as in equation (3.12) where $\tau = \frac{30}{365}$. Carr and Madan (1998), Britten-Jones and Neuberger (2000) and Jiang and Tian (2005) show that the VIX can be computed from the prices of call and put options with the same maturity at different strike prices:

$$VIX_t^2 = \frac{2e^{r_f \tau}}{\tau}\left[\int_0^{S_t^*} \frac{1}{K^2} P_t(K)dK + \int_{S_t^*}^{\infty} \frac{1}{K^2} C_t(K)dK\right], \tag{3.19}$$

where $K$ is the strike price, $C_t$ and $P_t$ are the put and call prices, $S_t^*$ is the fair forward price of the S&P500 index, and $r_f$ is the risk-free rate. The VIX index published by the CBOE is discretized, truncated, and interpolated across the two nearest maturities to achieve a constant 30-day maturity.[47] Jiang and Tian (2005) show through simulation analysis that the approximations used in the VIX index calculation are quite accurate.

Since February 2006, options on the VIX have been trading on the CBOE, which give investors a way to trade the volatility of volatility. As of Q3 2012, the open interest in front-month VIX options was about 2.5 million contracts, which is similar to the open interest in front-month S&P500 index option contracts.

[47] More details on the exact implementation of the VIX can be found in the white paper available on the CBOE website: http://www.cboe.com/micro/vix

We calculate our measure of the implied volatility of volatility using the same method as the VIX, applied to VIX options instead of S&P500 options. The index, which has since been published by the CBOE as the "VVIX index" in 2012 and back-filled, is calculated as:

$$VVIX_t^2 = \frac{2e^{r_f \tau}}{\tau} \left[ \int_0^{F_t} \frac{1}{K^2} P_t^*(K) dK + \int_{F_t}^{\infty} \frac{1}{K^2} C_t^*(K) dK \right], \tag{3.20}$$

where $F_t$ is the VIX futures price, and $C_t^*, P_t^*$ are the prices of call and put options on the VIX, respectively.[48] The squared VVIX is calculated from a portfolio of out-of-the-money call and put options on VIX futures contracts. It captures the implied volatility of VIX futures returns over the next 30-days, and is a model-free, forward-looking measure of the implied volatility of volatility.

In addition to the implied volatilities, we can also compute the realized volatilities for the stock market and the VIX. The construction here follows Barndorff-Nielsen and Shephard (2004) using high-frequency, intraday data.[49] Realized variance is defined as the sum of squared high-frequency log returns over the trading day:

$$RV_t = \sum_{j=1}^{N} r_{t,j}^2. \tag{3.21}$$

Barndorff-Nielsen and Shephard (2004) show that $RV_t$ converges to the quadratic variation as $N \to \infty$. We follow the standard approach of considering 5 minute return intervals. A finer sampling frequency results in better asymptotic properties of the realized variance estimator, but also introduces more market microstructure noise such as the bid-ask bounce discussed in Heston, Korajczyk, and Sadka (2010). Liu, Patton, and Sheppard (2013) show that the 5 minute realized variance is very accurate, difficult to beat in practice, and is

[48] The official index is back-filled until 2007. We apply the same methodology and construct the index for an additional year back to 2006. The correlation between our measure of the VVIX and the published index is over 99% in the post-2007 sample. Our empirical results remain essentially unchanged if we restrict our sample to only the post-2007 period.

[49] The data is obtained from http://www.tickdata.com.

typically the ideal sampling choice in most applications combining accuracy and parsimony.

We estimate two realized variance measures, one for the S&P500 and one for the VIX. For the S&P500, we use the S&P500 futures contract and the resulting realized variance will be denoted $RV^{SPX}$. For the VIX we use the spot VIX index and denote the resulting realized variance of the VIX by $RV^{VIX}$, which is our measure of the physical volatility of volatility.

*3.3.2. Variance Dynamics*

All of our variables are at the monthly frequency. The implied variance measures are given by the index values at the end of the month, and the realized variance measures are calculated over the past month and annualized.

Table 3.1 presents summary statistics for the implied and realized variance measures. While the average level of the VIX is about 24%, the average level of the VVIX is much higher at about 87%, which captures the fact that VIX futures returns are much more volatile than market returns: volatility, itself, is very volatile. $VVIX^2$ is also more volatile and less persistent than $VIX^2$, with an AR(1) coefficient of 0.423 compared to 0.805 for $VIX^2$. The VVIX exhibits relatively low correlation with the VIX, with a correlation coefficient of about 0.30. The mean of realized variance for S&P500 futures returns is 0.031, which corresponds to an annualized volatility of 17.6%. S&P500 realized variance is persistent and quite strongly correlated to the VIX index (correlation coefficient of 0.88) and much more weakly correlated to the VVIX index (correlation coefficient about 0.32). The realized variance of VIX is strongly related to the VVIX index (correlation of 0.53), and to a lesser extent, the VIX index (correlation of 0.38).

Figure 3.1 shows the time-series of the VIX and VVIX from February 2006 to June 2013. There are some common prominent moves in both series, such as the a peak during the financial crises. Notably, however, the VVIX also peaks during other times of economic uncertainty such as the summer of 2007 (quant meltdown, beginning of the subprime crisis), May 2010 (Eurozone debt crisis, flash crash) and August 2011 (U.S. debt ceiling crisis).

During these events, the VIX experienced upward movements, but of a magnitude far smaller than the spikes in the VVIX. The plot suggests that the VVIX captures important uncertainty-related risks in the aggregate market, distinct from the VIX itself.

In Figure 3.2 we present time-series plots for both S&P500 and VIX realized variances. As shown in Panel A, the realized and implied variances of the stock market follow a similar pattern, and S&P500 realized variance is nearly always below the implied variance. There is a large spike in both series around the financial crisis in October 2008, at which point realized variance exceeded implied variance. The difference between the mean of $VIX^2$ and $RV^{SPX}$ is typically interpreted as a variance premium, which is the difference between end-of-month model-free, forward-looking implied variance calculated from S&P500 index options and the realized variance of S&P500 futures returns over the past month. Unconditionally, the average level of the VIX (24%) is greater than the average level of the S&P500 realized volatility (17.6%), so that the variance premium is positive, consistent with the evidence in Bollerslev et al. (2009) and Drechsler and Yaron (2011). This implies that under the risk-neutral measure, volatility has a higher mean than under the physical probability measure. In turn, this evidence suggests that the market price of the volatility risk is negative.

Panel B of the Figure shows the time series of the realized and implied volatility of the VIX index. Generally, the implied volatility tends to increase at times of pronounced spikes in the realized volatility. The implied volatility is also high during other times of economic distress and uncertainty, such as May 2010 (Eurozone debt crisis and flash crash), and August 2011 (U.S. debt ceiling crisis). The VVIX largely follows the same pattern. During normal times, the VVIX is above the VIX realized variance, although during times of extreme distress we see the realized variance of VIX can exceed the VVIX. The average level of the VVIX (87%) is greater than the average level of the VIX realized volatility (73.4%), so that the volatility-of-volatility premium is also positive.[50] Similar to our discussion of the variance

[50]Song (2013) shows that the average level of his VVIX measure, computed using numerical integration rather than the model-free VIX construction, is lower than the average realized volatility of VIX. One of the key differences between his and our computations is the frequency of returns used in the realized variance computations. Consistent with the literature, we rely on 5-minute returns to compute the realized variances,

premium, this evidence suggests that investors dislike volatility-of-volatility risks, and the market price of volatility-of-volatility risks is negative.

In addition to unconditional moments, we can also analyze the conditional dependence of volatility and volatility of volatility. Specifically, we consider the predictability of future realized variances by the VIX and VVIX, in spirit of Canina and Figlewski (1993), Christensen and Prabhala (1998), and Jiang and Tian (2005) who use option implied volatilities to predict future realized volatilities. We follow Christensen and Prabhala (1998) and Jiang and Tian (2005) and conduct our predictability regressions of future realized variances using monthly, non-overlapping samples. We follow a standard approach in the literature and consider both univariate and multivariate encompassing regressions to assess the predictability of future realized variances by the VIX and VVIX.

In our main specification, the dependent variable is the realized variance ($RV$) over the next month, for both the S&P500 and the VIX. Univariate regressions test whether each implied volatility measure (the VIX or the VVIX) can forecast future realized variances; multivariate encompassing regressions compare the relative forecasting importance of the VIX and VVIX and whether one implied volatility measure subsumes the information content of the other. The univariate regressions are restricted versions of the corresponding multivariate encompassing regression, which are presented below:

$$RV_{t+1}^{SPX} = \beta_0 + \beta_1 VIX_t^2 + \beta_2 VVIX_t^2 + \beta_3 RV_t^{SPX} + \epsilon_{t+1}, \tag{3.22}$$

Similarly for the VIX, we have:

$$RV_{t+1}^{VIX} = \beta_0 + \beta_1 VIX_t^2 + \beta_2 VVIX_t^2 + \beta_3 RV_t^{VIX} + \epsilon_{t+1}. \tag{3.23}$$

Our benchmark results are presented for all variables calculated in annualized variance units.

while Song (2013) uses daily returns.

The first regression in Panel A of Table 3.2 shows that the VIX can forecast future realized variance of S&P500 returns. This is consistent with the findings of Jiang and Tian (2005). The VVIX can also forecast future S&P500 realized variance somewhat, although the statistical significance is weaker than that of the VIX and the magnitude of the regression coefficient is several times smaller. In the encompassing regression, we see that the VIX dominates the VVIX in forecasting future S&P500 realized variance. A one standard deviation increase in $VIX^2$ is associated with a 0.6 standard deviation increase in the realized variance of S&P500 returns next month. The coefficient on the VIX does not change much when we include the VVIX, which is consistent with our model. Including lags of the realized variances themselves do not materially change the results.

Panel B of Table 3.2 shows our predictability results for VIX realized variance, which is our proxy for physical volatility of volatility. The VIX is positively related to future VIX realized variation, but is not a significant predictor. The t-statistic is 0.95, and the adjusted $R^2$ is below zero. In stark contrast, the VVIX is a significant predictor of future VIX realized variation. The regression coefficient for the VVIX is about 0.8 in a univariate regression, and is largely unchanged in the multivariate regression. A one standard deviation increase in the current value of $VVIX^2$ is associated with more than one-third standard deviation increase in next month realized variance of VIX.

The empirical evidence suggests that fluctuations in the volatility of volatility are not directly related to the level of the volatility itself. This is consistent with our two-volatility model specification in Section 2. In many reduced form and structural models, the volatility of volatility is directly linked to the level of the volatility. For example, Heston (1993) models volatility as following a Cox et al. (1985) square-root process. In that case, the level of volatility itself should forecast future realized volatility of volatility. The evidence in the data does not support this assumption, and calls for a richer dynamics of the volatility process, with separate movements in the volatility of volatility.

## 3.4. Evidence from Options

In this section, we analyze the implications of equity and VIX option price dynamics for the pricing of volatility and volatility-of-volatility risks in the data. Our economic model suggests that the market prices of volatility and volatility-of-volatility risks determine the key properties of the cross-section and time-series of delta-hedged equity and VIX option gains. Specifically, if market prices of volatility and volatility risks are negative, the average delta-hedged equity and VIX option gains are also negative. In the cross-section, the average returns are more negative to the option strategies which have higher exposure to the volatility and volatility-of-volatility risks. Finally, in the time series higher volatility and volatility of volatility predicts more negative gains in the future. We evaluate these model predictions in the data, and find a strong support that both volatility and volatility-of-volatility risks are priced in the option markets, and have negative market prices of risks.

### *3.4.1. Delta-Hedged Option Gains*

We consider discrete-time counterparts to the continuously-rebalanced delta-hedged gains in equations (3.7) and (3.17):

$$\begin{aligned}
\Pi_{t,t+\tau} &= \underbrace{C_{t+\tau} - C_t}_{\text{option gain/loss}} - \underbrace{\sum_{n=0}^{N-1} \Delta_{t_n} \left(S_{t_{n+1}} - S_{t_n}\right)}_{\text{delta hedging gain/loss}} + \underbrace{\sum_{n=0}^{N-1} r_f \left(\Delta_{t_n} S_{t_n} - C_t\right) \frac{\tau}{N}}_{\text{risk-free rate}}, \\
\Pi^*_{t,t+\tau} &= \underbrace{C^*_{t+\tau} - C^*_t}_{\text{option gain/loss}} - \underbrace{\sum_{n=0}^{N-1} \Delta_{t_n} \left(F_{t_{n+1}} - F_{t_n}\right)}_{\text{delta hedging gain/loss}} - \underbrace{\sum_{n=0}^{N-1} r_f C^*_t \frac{\tau}{N}}_{\text{risk-free rate}}.
\end{aligned} \tag{3.24}$$

$\Delta_{t_n}$ indicates option delta, e.g. $\Delta_{t_n} = \frac{\partial C_{t_n}}{\partial S_{t_n}}$, and $N$ is the number of trading days in the month. This discrete delta-hedging scheme is also used in Bakshi and Kapadia (2003) and Bertsimas et al. (2000).

At the close of each option expiration, we look at the prices of all options with non-zero open interest and non-zero trading volume. We take a long position in the option, and

hedge the $\Delta$ each day according to the Black-Scholes model and hedge the $\Delta$ risk, with the net investment earning the risk-free interest rate appropriately.[51] To minimize the effect of recording errors, we discard options that have implied volatilities below the 1st percentile or above the 99th percentile. All options have exactly one calendar month to maturity; S&P500 options expire on the third Friday of every month, while VIX options expire on the Wednesday that is 30 days away from the third Friday of the following month.

Table 3.3 shows average index and VIX delta-hedged option gains in our sample. We separate options by call or put, and group each option into four bins by moneyness to obtain eight bins for both S&P500 and VIX options. The first column $\frac{\Pi}{S}$ gives the delta-hedged option gain scaled by the index level, and the second column $\frac{\Pi}{C}$ gives the delta-hedged option gain scaled by the option price, which can be interpreted more readily as a "return" in the traditional sense. Panel A of Table 3.3 show that the average out-of-the-money delta-hedged S&P500 call options have significantly negative returns. Likewise, delta-hedged put options on the S&P500 also have significantly negative returns at all levels of moneyness. This evidence is largely consistent with Bakshi and Kapadia (2003), who focus on call options in an earlier sample period. In the model, negative average returns on delta-hedged index calls imply that volatility and/or volatility-of-volatility risks have a negative market price of risk. S&P500 option gains display mild positive serial correlation, which we will account for in our time-series predictive regressions in the later sections.

Panel B of Table 3.3 shows the average returns for delta-hedged VIX options. The average delta-hedged VIX option returns are negative and statistically significant in all bins except for out-of-the-money puts, which are marginally significant. Call options lose more money as they become more out of the money, regardless of whether we are scaling by the index or by the option price. Estimates of the loss for call options ranges from -0.57% of the

[51] This requires an estimate of the implied volatility of the option, which may require an option price. We use implied volatilities directly backed out from market prices of options whenever possible; if an option does not have a quoted price on any intermediate date, we fit a cubic polynomial to the implied volatility curve given by options with quoted prices, and back out the option's implied volatility. This is similar to typical option position risk management done by professional traders.

index value for in-the-money VIX calls to -1.41% of the index value for out-of-the-money VIX calls. When viewed as a percentage of the option price, at-the-money delta-hedged VIX calls return about -10% per month. The results for VIX put options are similar. In the model, negative average returns for delta-hedged VIX options imply that investors dislike volatility-of-volatility risks, and are systematically paying a premium hedge against increases in the volatility of volatility. This suggests that the price of the volatility-of-volatility risk is negative. VIX option gains have small negative serial correlation, which we also account for in our time-series predictive regressions in the later sections. For both the S&P500 and VIX, the delta-hedged option gains are quite volatile.

In the next section, we provide further direct evidence by controlling for the exposures of the delta-hedged option positions to the underlying risks.

*3.4.2. Cross-Sectional Evidence*

As shown in the previous section, the average delta-hedged option gains are negative for S&P500 and VIX options. Our model further implies (see equations (3.10) and (3.18)) that options with higher sensitivity (higher $\beta_t^V, \beta_t^\eta, \beta_t^*$) to volatility and volatility-of-volatility risks should have more negative gains. To compute the estimates of option exposures to the underlying risks, we follow the approach of Bakshi and Kapadia (2003) which relies on using the Black and Scholes (1973b) model to proxy for the true option betas.

Specifically, to compute the proxy for the option beta to volatility risk, we consider the vega of the option:

$$\frac{\partial C}{\partial \sigma} = S\sqrt{\frac{\tau}{2\pi}}e^{-\frac{d_1^2}{2}} \propto e^{-\frac{d_1^2}{2}}, \tag{3.25}$$

where $d_1 = \frac{1}{\sigma\sqrt{\tau}}\left[\log\frac{S}{K} + \left(r_f - q + \frac{\sigma^2}{2}\right)\tau\right]$, $q$ is the dividend yield, and $\sigma$ is the implied volatility of the option. This approach allows us to compute proxies for the exposures of equity options to volatility risks, and of VIX options to the volatility-of-volatility risks.

To illustrate the relation between the moneyness of the option and the vega-measured exposure of options to volatility risks, we show the option vega as a function of option moneyness in Figure 3.3. Vega represents an increase in the value of the option as implied volatility increases by 1%. Higher volatility translates into higher future profits from delta-hedging due to the convexity effect; hence both call and put options have strictly positive vegas. Further, as the curvature of option value is the highest for at-the-money options, at-the-money options have the highest vega in the cross-section, and thus the largest exposure to the volatility risks. An alternative way to proxy for the option sensitivity to volatility risks is to use "gamma" of the option which represents the second derivative of the option price to the underlying stock price: $\Gamma = \frac{\partial^2 C}{\partial S^2}$. As shown in Figure 3.3, the shape of the vega and gamma functions are almost identical, hence, the implied cross-sectional dispersion in volatility betas by moneyness are very similar as well.

To capture the sensitivity of option price to the volatility of volatility, we compute the Black and Scholes (1973b) second partial derivative of the option price with respect to the volatility, which is known in "volga" for "volatility gamma". Volga is calculated as:

$$\frac{\partial^2 C}{\partial \sigma^2} = S\sqrt{\frac{\tau}{2\pi}} e^{-\frac{d_1^2}{2}} \left(\frac{d_1 d_2}{\sigma}\right) = \frac{\partial C}{\partial \sigma}\left(\frac{d_1 d_2}{\sigma}\right), \tag{3.26}$$

$d_2 = d_1 - \sigma\sqrt{\tau}$. Figure 3.4 shows the plot of volga as a function of the moneyness of the option. Volga is positive, and exhibits twin peaks with a valley around at-the-money. At-the-money options are essentially pure bets on volatility, and are approximately linear in volatility (see Stein (1989)). Therefore, the volga is the lowest for at-the-money options. Deep-out-of-the-money options and deep-in-the-money options do not have much sensitivity to volatility of volatility either, since for the former it is a pure directional bet, and for the latter the option value is almost entirely comprised of intrinsic value. Options that are somewhat away from at-the-money are most exposed to volatility-of-volatility risks.

Table 3.4 shows our cross-sectional evidence from the regressions of average option returns on our proxies of options' volatility and volatility-of-volatility betas. Panel A shows uni-

variate and multivariate regressions of delta-hedged S&P500 option gains scaled by the index on the sensitivities of the options to volatility and volatility-of-volatility risks. Our encompassing regression for delta-hedged S&P500 options is:

$$GAINS^i_{t,t+\tau} = \frac{\Pi^i_{t,t+\tau}}{S_t} = \tilde{\lambda}_1 VEGA^i_t + \tilde{\lambda}_2 VOLGA^i_t + \gamma_t + \epsilon^i_{t,t+\tau}. \tag{3.27}$$

Since each date includes multiple options, as in Bakshi and Kapadia (2003) we allow for a date-specific component in $\Pi^i_{t,t+1}$ due to the option expirations. Conceptually, our approach is related to Fama and MacBeth (1973) regressions. Instead of estimating risk betas in the first stage, due to the non-linear structure of option returns, we measure the our exposures from economically motivated proxies for the risk sensitivities.[52]

The results in Panel A show that both volatility and the volatility of volatility are priced in the cross-section of delta-hedged S&P500 option returns. Options more exposed to volatility and volatility-of-volatility risks have more negative expected returns. The univariate estimates for vega and volga are -0.051 and -0.007. Both t-statistics are highly significant at conventional levels. In the multivariate regression, we see that the statistical significance becomes stronger for both risks, and the point estimates become larger: -0.178 for vega and -0.019 for volga. The signs and significance of these coefficients implies a significant negative volatility-of-volatility risk premium.

Panel B of Table 3.4 presents cross-sectional results for delta-hedged VIX options. As equation (3.18) shows, delta-hedged VIX options are no longer exposed to volatility risk, and the vega for VIX options captures the sensitivity of VIX options to innovations in the volatility of volatility. The coefficient on vega of -1.46 is negative and statistically significant. Thus, in the cross-section of both S&P500 options and VIX options, we find strong evidence of a negative price of volatility-of-volatility risk.

[52]Song and Xiu (2013) demonstrate an alternative method of estimating risk sensitivities nonparametrically using local linear regression methods.

*3.4.3. Time-Series Evidence*

In the model, time-variation in the expected delta-hedged option gains is driven by $V_t$ and $\eta_t$, and the loadings are determined by the market prices of volatility and volatility-of-volatility risk. We group options into the same bins as we used for average returns in Table 3.3, and average the scaled gains within each bin, so that we have a time-series of option returns for each moneyness bin. To examine the contribution of both risks for the time-variation in expected index option payoffs, we consider the following regression:

$$GAINS^i_{t,t+\tau} = \frac{\Pi^i_{t,t+\tau}}{S_t} = \beta_0 + \beta_1 VIX^2_t + \beta_2 VVIX^2_t + \gamma GAINS^i_{t-\tau} + u^i + \epsilon^i_{t+\tau}. \quad (3.28)$$

where we include fixed effects $u^i$ to account for the heterogeneity in the sensitivity of options in different moneyness bins to the underlying risks. We regress the delta-hedged option gain scaled by the index from expiration to expiration on the value of the VIX and VVIX indices at the end of the earlier expiration; in other words, we run one-month ahead predictive regressions of delta-hedged option returns on the VIX and VVIX. We include lagged gains to adjust for serial correlation in the residuals, following Bakshi and Kapadia (2003).

Panel A of Table 3.5 shows the regression results for the index options. The univariate regression of delta-hedged S&P500 option gains on $VIX^2$ is negative and statistically significant, which is consistent with Bakshi and Kapadia (2003). The second regression is a multivariate regression of delta-hedged option gains on $VVIX^2$, which shows that both the VIX and VVIX loadings are negative and statistically significant. A one standard deviation increase in $VIX^2$ is associated with a -0.047% (of the S&P500 index value) lower delta-hedged option gain. In the same regression, a one standard deviation increase in $VVIX^2$ is associated with a -0.10% (of the S&P500 index value) lower delta-hedged option gain. Hence, both volatility and volatility of volatility command negative prices of risk in the S&P500 options market, and have significant contribution to the fluctuations in expected option returns.

Panel B of Table 3.5 shows the corresponding evidence for VIX options. The VVIX negatively and significantly predicts future VIX option gains. This is consistent with a negative market price of volatility-of-volatility risk.

## 3.5. Robustness

### *3.5.1. Alternative Variance Specifications*

Our results for the predictability of realized by implied variance are robust to alternative specifications of volatility. Specifically, we consider regressing in volatility units or log-volatility units, rather than variance. The robustness regressions follow the form:

$$\sigma^x_{t+1} = \beta_0 + \beta_1 VIX_t + \beta_2 VVIX_t + \epsilon_{t+1} \tag{3.29}$$

$$\ln \sigma^x_{t+1} = \beta_0 + \beta_1 \ln VIX_t + \beta_2 \ln VVIX_t + \epsilon_{t+1} \tag{3.30}$$

where $x$ refers to $SPX$ or $VIX$.

In Table 3.6, we see that the point estimates and significance are very close to our baseline specification in variance units. In fact, the log-volatility specifications have coefficients much closer to 1 for the S&P500 predictability results, suggesting that the VIX is generally an unbiased forecast of future realized S&P500 volatility over the next month.

### *3.5.2. Sensitivity to Jump Risk Measures*

The evidence in our paper highlights the roles of the volatility and volatility-of-volatility factors, which are driven by smooth Brownian motion shocks. In principle, the losses on delta-hedged option portfolios can also be attributed to large, discontinuous movements (jumps) in the stock market and in the market volatility. In this Section we verify that our empirical evidence for the importance of the volatility-related factors is robust to the inclusion of jump measures considered in the literature.

Specifically, we consider three measures of jump risks, which we construct for the S&P500

returns and the VIX. Our first jump measure corresponds to the slope of the implied volatility curve:

$$\begin{aligned} SLOPE^{SPX} &= \sigma_{OTM}^{SPX} - \sigma_{ATM}^{SPX}, \\ SLOPE^{VIX} &= \sigma_{OTM}^{VIX} - \sigma_{ATM}^{VIX}. \end{aligned} \tag{3.31}$$

The $OTM$ contract for the S&P500 options is defined as a put option with a moneyness closest to 0.9, and for VIX options as a call option with a moneyness closest to 1.1. In both cases, the $ATM$ option has moneyness of 1. These slopes are positive for both index and VIX options. Positive slope of the index volatility smile is consistent with the notion of negative jumps in market returns (see e.g. Bates (2000), Pan (2002), Eraker et al. (2003)), while the fact that the implied volatility curve for VIX options slopes upwards (call options are more expensive than put options on average) is consistent with the positive volatility jumps (Drechsler and Yaron (2011) and Eraker and Shaliastovich (2008), among others). In this sense, these slope measures help capture the variation in the market and volatility jumps in the economy.

Our second jump measure incorporates the whole cross-section of option prices, beyond just the slope of the smile. It is based on the model-free risk-neutral skewness of Bakshi et al. (2003):

$$SKEW(t, t+\tau) = \frac{e^{r_f \tau} W_{t,t+\tau} - 3\mu_{t,t+\tau} e^{r_f \tau} V_{t,t+\tau} + 2\mu_{t,t+\tau}^3}{\left[e^{r_f \tau} V_{t,t+\tau} - \mu_{t,t+\tau}^2\right]^{3/2}}, \tag{3.32}$$

where $V_{t,t+\tau}, W_{t,t+\tau}, X_{t,t+\tau}$ are given by the prices of the volatility, cubic, and quartic contracts. Importantly, these measures are computed model-free using the observed option prices. The details for the computations are provided in the Appendix.

Finally, our third measure of jump risks is based on the high-frequency index and VIX data, rather than the option prices. It corresponds to the realized jump intensity, and relies on the bipower variation methods in Barndorff-Nielsen and Shephard (2004), Huang and Tauchen

(2005), and Wright and Zhou (2009). Specifically, while the realized variance defined in (3.21) captures both the continuous and jump variation, the bipower variation, defined as:

$$BV_t = \frac{\pi}{2}\left(\frac{M}{M-1}\right)\sum_{j=2}^{M}|r_{t,j-1}||r_{t,j}| \tag{3.33}$$

measures the amount of continuous variation returns. Hence, we can use the test statistic to determine if there is a jump on any given day:

$$J_t = \frac{\frac{RV_t - BV_t}{RV_t}}{\sqrt{\frac{\theta}{M}\max 1, \frac{QV_t}{BV_t^2}}}, \tag{3.34}$$

where $\theta = \left(\frac{\Pi}{2}\right)^2 + \pi - 5$, and $QV_t$ is the quad-power quarticity defined in Huang and Tauchen (2005) and Barndorff-Nielsen and Shephard (2004). The test statistic is distributed as $\mathcal{N}(0,1)$. We flag the day as having a jump if the probability exceeds 99.9% both for index returns and for the VIX. These cut-offs imply an average frequency of jumps of once every two months for the index, and about three jumps a month for the VIX. This is broadly consistent with the findings of Tauchen and Todorov (2011), who find that VIX jumps tend to happen much more frequently than S&P500 jumps. Over a month, we sum up all the days where we have a jump, and we define our jump intensity measure on a monthly level as:

$$RJ = \frac{1}{T}\sum_{i=0}^{T-1} J_{t+i},$$

where $T$ is the number of trading days in the month.

We use the jump statistics to document the robustness of the link between the volatility and volatility-of-volatility factors and options gains. We consider a regression:

$$GAINS^i_{t,t+\tau} = \beta_0 + \beta_1 VIX_t^2 + \beta_2 VVIX_t^2 + \beta_3 JUMP_t + \gamma GAINS^i_{t-\tau} + u^i + \epsilon^i_{t+\tau} \tag{3.35}$$

where $JUMP_t$ is one of the above jump risk proxies. We use index jump measures for

index gains, and VIX jump measures for VIX gains. Table 3.7 displays our results. Both for S&P500 and VIX options, controlling for $SLOPE$ does not change the ability of $VIX$ and $VVIX$ to predict future delta-hedged option gains. Both factors are still significant, and the point estimates $\hat{\beta}_1$ and $\hat{\beta}_2$ are largely unchanged. $SLOPE$ itself is not significant at conventional levels for S&P500 options or VIX options. When we control for realized jump intensity $RJ$, we see a similar result where the statistical significance of $VIX$ and $VVIX$ as well as their point estimates are largely unchanged. Neither for S&P500 nor for VIX options, $RJ$ does not seem to be a significant predictor of future delta-hedged option gains. For VIX options, $RJ$ does not affect the point estimate on $VVIX$ nor its significance; however, $RJ$ does seem to be a predictor of future VIX option gains. Finally, risk-neutral skewness also does not affect the predictive ability of $VIX$ and $VVIX$. While the skewness measures as insignificant themselves, the estimates have the correct sign since skewness is negative for S&P500 options and positive for VIX options; this is broadly similar to the findings of Bakshi and Kapadia (2003).

Hence, our evidence suggests that the VIX and VVIX have a significant impact on option returns even in the presence of stock market and volatility jumps. We leave a formal treatment of jumps for future research.

## 3.6. Conclusion

Using S&P500 and VIX options data, we show that a time-varying volatility of volatility is a separate risk factor which affects the option returns, above and beyond volatility risks. We measure volatility risks using the VIX index, and volatility-of-volatility risk using the VVIX index. The two indices, constructed from the index and VIX option data, capture the ex-ante risk-neutral uncertainty of investors about future market returns and VIX innovations, respectively. The VIX and VVIX have separate dynamics, and are only weakly related in the data: the correlation between the two series is 0.30. On average, risk-neutral volatilities identified by the VIX and VVIX exceed the realized physical volatilities of the corresponding variables in the data. Hence, the variance premium and variance-of-variance premium for

VIX are positive, which suggests that investors dislike variance and variance-of-variance risks.

We show the pricing implications of volatility and volatility-of-volatility risks using options market data. Average delta-hedged option gains are negative, which suggests that investors pay a premium to hedge against innovations in not only volatility but also the volatility of volatility. In the cross-section of both delta-hedged S&P500 options and VIX options, options with higher sensitivities to volatility-of-volatility risk earn more negative returns. In the time-series, higher values of the VVIX predict more negative delta-hedged option returns, for both S&P500 and VIX options.

Our findings are consistent with a no-arbitrage model which features time-varying market volatility and volatility-of-volatility factors. The volatility factors are priced by the investors, and in particular, volatility and volatility of volatility have negative market prices of risks.

## Tables

Table 3.1: Summary Statistics

| Variable | Mean | Std. Dev. | AR(1) | Corr. $VIX^2$ | Corr. $VVIX^2$ |
|---|---|---|---|---|---|
| $VIX^2$ | 0.059 | 0.060 | 0.805 | 1.000 | 0.301 |
| $VVIX^2$ | 0.763 | 0.197 | 0.423 | 0.301 | 1.000 |
| $RV^{SPX}$ | 0.031 | 0.060 | 0.620 | 0.880 | 0.316 |
| $RV^{VIX}$ | 0.539 | 0.434 | 0.192 | 0.378 | 0.526 |

Table 3.1 gives summary statistics for the variance and jump measures we use. $RV^{SPX}$ is the realized variance of S&P500 returns, calculated using 5-minute log futures returns. Monthly variables from 2006m2 to 2013m7. $RV^{VIX}$ is the realized variance of VIX innovations, calculated using 5-minute log VIX index innovations. Realized measures are annualized, and overnight returns are excluded. AR(1) is the persistence of the variable. Corr. to $VIX^2$ is the correlation of the variable with $VIX^2$, and Corr. to $VVIX^2$ is similarly defined. The model-free implied variance variable $VIX^2$ is defined as $\left(\frac{VIX}{100}\right)^2$ and $VVIX^2$ is defined as $\left(\frac{VVIX}{100}\right)^2$.

Table 3.2: Predictability of Realized Measures

| | $VIX^2$ | | $VVIX^2$ | | $R^2_{adj}$ |
|---|---|---|---|---|---|
| | Slope | T-stat. | Slope | T-stat. | |
| **Panel A: S&P500 Index** | | | | | |
| $RV^{SPX}_{t,t+1}$ | 0.611 | [4.97] | | | 37.46 |
| | | | 0.066 | [1.57] | 3.60 |
| | 0.601 | [4.62] | 0.010 | [0.55] | 36.83 |
| **Panel B: VIX Index** | | | | | |
| $RV^{VIX}_{t,t+1}$ | 0.610 | [0.95] | | | -0.43 |
| | | | 0.799 | [4.23] | 12.20 |
| | -0.192 | [-0.30] | 0.817 | [3.52] | 11.23 |

Table 3.2 gives the realized measure predictability regressions. Monthly frequency sample spanning 2006m2 to 2013m6. The t-statistics shown are calculated using Newey and West (1987) robust standard errors with 6 lags. Realized measures calculated using high-frequency 5-minute data. Numbers in brackets are t-statistics.

Table 3.3: Delta-Hedged Option Gains

| | Moneyness ($\frac{Strike}{Forward}$) | $\frac{\Pi}{S}$(%) | t-stat. | Std. Dev. | AR(1) | $\frac{\Pi}{C}$(%) | t-stat. |
|---|---|---|---|---|---|---|---|
| **Panel A: S&P500 Index** | | | | | | | |
| Call Options | 0.950 to 0.975 | 0.08 | [ 2.20] | 0.71 | 0.41 | 1.75 | [ 2.37] |
| | 0.975 to 1.000 | 0.01 | [ 0.50] | 0.60 | 0.26 | 0.76 | [ 0.82] |
| | 1.000 to 1.025 | -0.08 | [-3.32] | 0.54 | 0.11 | -5.87 | [-3.44] |
| | 1.025 to 1.050 | -0.13 | [-5.89] | 0.48 | 0.13 | -32.43 | [-7.18] |
| Put Options | 0.950 to 0.975 | -0.16 | [-4.84] | 0.72 | 0.38 | -18.27 | [-5.41] |
| | 0.975 to 1.000 | -0.21 | [-7.39] | 0.62 | 0.24 | -11.47 | [-6.48] |
| | 1.000 to 1.025 | -0.27 | [-9.79] | 0.59 | 0.05 | -9.81 | [-10.68] |
| | 1.025 to 1.050 | -0.30 | [-9.85] | 0.54 | 0.07 | -6.74 | [-10.23] |
| **Panel B: VIX Index** | | | | | | | |
| Call Options | 0.800 to 0.900 | -0.57 | [-3.07] | 2.07 | -0.09 | -3.44 | [-3.23] |
| | 0.900 to 1.000 | -1.35 | [-5.75] | 2.55 | -0.21 | -11.88 | [-5.73] |
| | 1.000 to 1.100 | -1.08 | [-3.72] | 2.86 | -0.09 | -12.56 | [-3.27] |
| | 1.100 to 1.200 | -1.41 | [-5.11] | 2.74 | -0.09 | -22.07 | [-4.48] |
| Put Options | 0.800 to 0.900 | -0.57 | [-1.79] | 2.49 | -0.06 | -13.32 | [-1.43] |
| | 0.900 to 1.000 | -1.28 | [-5.14] | 2.67 | -0.19 | -17.71 | [-4.27] |
| | 1.000 to 1.100 | -1.04 | [-3.39] | 2.99 | -0.12 | -7.50 | [-3.04] |
| | 1.100 to 1.200 | -1.30 | [-4.61] | 2.74 | -0.11 | -6.19 | [-4.40] |

Table 3.3 gives the delta-hedged option scaled gains. Monthly frequency sample spanning 2006m3 to 2012m10 for S&P500 options and VIX options. The t-statistics test the null hypothesis that the delta-hedged option scaled gain is equal to zero. Options have one month to maturity, and are held expiration to expiration. $\Pi$ is the delta-hedged option gain, as given in Bakshi and Kapadia (2003). Delta values calculated using the Black-Scholes formula. On intermediate dates when an option price cannot be found, we fit a cubic polynomial to observed market implied volatilities to back out the option's implied volatility, which we use to calculate a delta for hedging purposes. The portfolio gains can be interpreted as an equal-weighted portfolio of all options whose moneyness falls inside the corresponding bin. The delta-hedge is rebalanced daily, with the margin difference earning the risk-free rate. $\frac{\Pi}{S}$ is the delta-hedged option gain scaled by the index, and $\frac{\Pi}{C}$ is the delta-hedged option gain scaled by the option price (for both puts and calls). Numbers in brackets are t-statistics that test the null hypothesis that the average delta-hedged gain is equal to zero.

Table 3.4: Delta-Hedged Option Gains by Volatility Risk Sensitivities

| | Vega $= \frac{\partial C}{\partial \sigma}$ | | Volga $= \frac{\partial^2 C}{\partial \sigma^2}$ | |
|---|---|---|---|---|
| | slope | t-stat. | slope | t-stat. |
| **Panel A: SPX Options** | | | | |
| $\frac{\Pi_{t,t+1}}{S_t}$ | -0.051 | [-2.24] | | |
| | | | -0.007 | [-3.29] |
| | -0.178 | [-8.09] | -0.019 | [-8.17] |
| **Panel B: VIX Options** | | | | |
| $\frac{\Pi_{t,t+1}}{S_t}$ | -1.68 | [-4.12] | | |

Table 3.4 gives the delta-hedged S&P500 and VIX option gains cross-sectional regressions. Cross-sectional regression at monthly frequency sample spanning 2006m3 to 2012m10, with time fixed effects, using robust standard errors. Constant is omitted because zero risk sensitivity implies zero expected delta-hedged option gain. The dependent variable is delta-hedged option gain $\Pi_{t,t+\tau}$, which is calculated as described in the data section, and scaled gains are given in percentages. Panel A is for S&P500 options. The independent variables are vega and volga, as calculated from Black-Scholes. Panel B is for VIX options. The independent variable in Panel B is vega (of VIX) calculated from Black-Scholes. Numbers in brackets are t-statistics.

Table 3.5: Predictability of Delta-Hedged SPX Option Gains

| | $VIX^2$ | | $VVIX^2$ | | $GAINS_{t-1}$ | |
|---|---|---|---|---|---|---|
| | slope | t-stat. | slope | t-stat. | slope | t-stat. |
| **Panel A: SPX Options** | | | | | | |
| $GAINS_t = \frac{\Pi_{t,t+\tau}}{S_t}$ | -0.78 | [-2.83] | | | 0.29 | [4.17] |
| | | | -0.51 | [-4.64] | 0.31 | [5.18] |
| | -0.47 | [-2.30] | -0.49 | [-4.85] | 0.34 | [4.79] |
| **Panel B: VIX Options** | | | | | | |
| $GAINS_t = \frac{\Pi_{t,t+\tau}}{S_t}$ | | | -1.37 | [-3.57] | -0.07 | [-3.40] |

Table 3.5 gives the delta-hedged S&P500 option gains (Panel A) and VIX option gains (Panel B) predictability regressions. Panel regression at monthly frequency sample spanning 2006m3 to 2012m10, with cross-sectional fixed effects, using robust standard errors. The dependent variable is delta-hedged option gain $\Pi_{t,t+\tau}$, which is calculated as described in the data section, and scaled gains are given in percentages. The cross-sectional identifier is the moneyness bin of the option, as given in Table 3.3. Gains for each bin are averaged, which can be interpreted as the gain on an equally-weighted portfolio of options within a given moneyness bin. The independent variables are $VIX^2 = \left(\frac{VIX}{100}\right)^2$, $VVIX^2 = \left(\frac{VVIX}{100}\right)^2$ which are in annualized percentage squared units. Following Bakshi and Kapadia (2003), lagged gains are included to correct for serial correlation of the residuals. Numbers in brackets are t-statistics.

Table 3.6: Predictability of Realized Measures - Alternate Specifications

| | $VIX$ | | $VVIX$ | | $\ln VIX$ | | $\ln VVIX$ | |
|---|---|---|---|---|---|---|---|---|
| | slope | t-stat. | slope | t-stat. | slope | t-stat. | slope | t-stat. |
| **Panel A: S&P500 Index** | | | | | | | | |
| $\sigma^{SPX}_{t,t+1}$ | 0.736 | [6.50] | | | | | | |
| | | | 0.179 | [1.51] | | | | |
| | 0.732 | [6.04] | 0.011 | [0.19] | | | | |
| $\ln \sigma^{SPX}_{t,t+1}$ | | | | | 0.961 | [8.72] | | |
| | | | | | | | 0.514 | [1.01] |
| | | | | | 0.969 | [8.07] | -0.110 | [-0.40] |
| **Panel B: VIX Index** | | | | | | | | |
| $\sigma^{VIX}_{t,t+1}$ | 0.210 | [0.83] | | | | | | |
| | | | 0.748 | [4.23] | | | | |
| | -0.014 | [-0.06] | 0.751 | [3.50] | | | | |
| $\ln \sigma^{VIX}_{t,t+1}$ | | | | | 0.047 | [0.55] | | |
| | | | | | | | 0.735 | [3.62] |
| | | | | | -0.006 | [-0.08] | 0.739 | [3.14] |

Table 3.6 gives the realized measure predictability regressions for alternative volatility specifications (volatility and log-volatility). Monthly frequency sample spanning 2006m2 to 2013m6. The t-statistics shown are calculated using Newey and West (1987) robust standard errors with 6 lags. Realized measures calculated using high-frequency 5-minute data. VIX and VVIX are in quoted index units divided by 100, and can be interpreted as percentage annual volatility. Numbers in brackets are t-statistics.

Table 3.7: Robustness to Jump Measures

| | $VIX^2$ | | $VVIX^2$ | | $SLOPE$ | | $SKEW$ | | $RJ$ | | $GAINS_{t-1}$ | |
|---|---|---|---|---|---|---|---|---|---|---|---|---|
| | slope | t-stat. | slope | t-stat. | slope | t-stat. | slope | t-stat. | slope | t-stat. | slope | t-stat. |
| **Panel A: SPX Options** | | | | | | | | | | | | |
| $GAINS_t = \frac{\Pi_{t,t+\tau}}{S_t}$ | -0.63 | [-2.12] | -0.42 | [-7.09] | -2.25 | [-1.54] | | | | | 0.33 | [4.91] |
| | -0.51 | [-2.13] | -0.48 | [-5.10] | | | 0.01 | [0.96] | | | 0.34 | [4.70] |
| | -0.47 | [-2.28] | -0.48 | [-5.30] | | | | | 0.01 | [0.71] | 0.34 | [4.74] |
| **Panel B: VIX Options** | | | | | | | | | | | | |
| $GAINS_t = \frac{\Pi_{t,t+\tau}}{S_t}$ | | | -1.36 | [-3.48] | -0.69 | [-0.36] | | | | | -0.07 | [-3.27] |
| | | | -1.37 | [-3.55] | | | -0.08 | [-0.88] | | | -0.08 | [-3.09] |
| | | | -1.37 | [-3.56] | | | | | -0.01 | [-0.51] | -0.07 | [-3.40] |

Table 3.7 gives the delta-hedged S&P500 option gains (Panel A) and VIX option gains (Panel B) predictability regressions. Panel regression at monthly frequency sample spanning 2006m3 to 2012m10, with cross-sectional fixed effects, using robust standard errors . The dependent variable is delta-hedged option gain $\Pi_{t,t+\tau}$, which is calculated as described in the data section, and scaled gains are given in percentages. The cross-sectional identifier is the moneyness bin of the option, as given in Table 3.3. Gains for each bin are averaged, which can be interpreted as the gain on an equally-weighted portfolio of options within a given moneyness bin. The independent variables are $VIX^2 = \left(\frac{VIX}{100}\right)^2$, $VVIX^2 = \left(\frac{VVIX}{100}\right)^2$ which are in annualized percentage squared units. For S&P500 options, $SLOPE$ is calculated as the Black and Scholes (1973b) implied volatility of an out-of-the-money ($\frac{K}{S} = 0.9$) minus the implied volatility of an at-the-money put option ($\frac{K}{S} = 1$). For VIX options, $SLOPE$ is calculated using the difference in implied volatility between a $\frac{K}{S} = 1.1$ call option and an at-the-money call option. $RJ$ is calculated using high-frequency data as in Barndorff-Nielsen and Shephard (2004) and Wright and Zhou (2009); the calculation for S&P500 options uses S&P500 futures tick data, while the calculation for VIX options uses VIX tick data. $SKEW$ is calculated using the model-free method of Bakshi et al. (2003). Following Bakshi and Kapadia (2003), lagged gains are included to correct for serial correlation of the residuals. Numbers in brackets are t-statistics.

## Figures

Figure 3.1: Time Series Plot

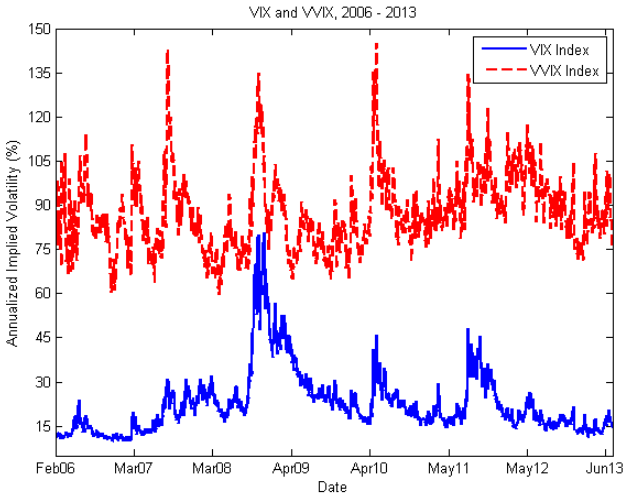



Figure 3.1 plots VIX and VVIX index from 2006m2 to 2013m6. The solid blue line is VIX and the dashed red line is VVIX. VVIX data from 2007m1 onwards is from CBOE, and for 2006 it is calculated from VIX options using the method published by the CBOE. All VIX data is from official CBOE VIX levels.

Figure 3.2: Realized Measures

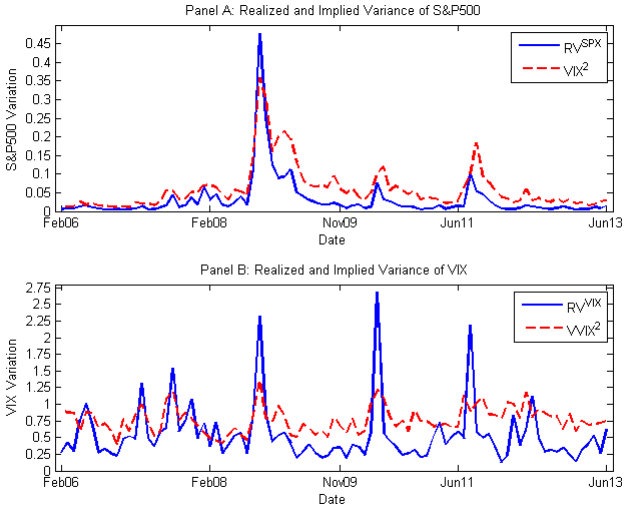



Figure 3.2 plots realized and implied variances for the S&P500 (top panel) and VIX (bottom panel). Monthly data from 2006m2 to 2013m6. The blue solid lines are realized variances, and red dashed lines are the model-free implied variances. Realized variances calculated from 5-minute high-frequency data. All measures in annualized variance units.

Figure 3.3: Vega and Gamma by Moneyness

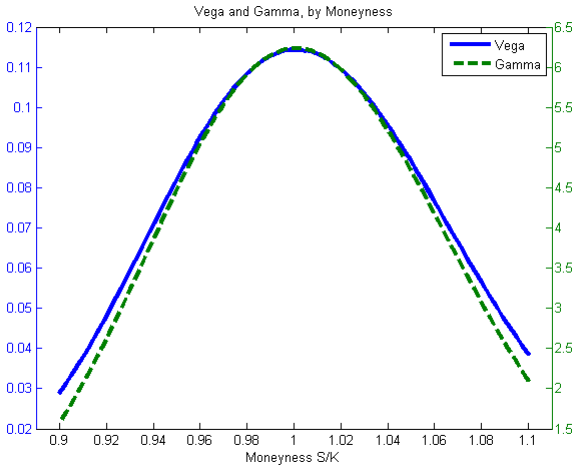



Figure 3.3 plots Black-Scholes vega and gamma by option moneyness for average levels of volatility. Although levels are different, the shape (hence cross-sectional dispersion) of the variables are nearly identical.

Figure 3.4: Volga by Moneyness

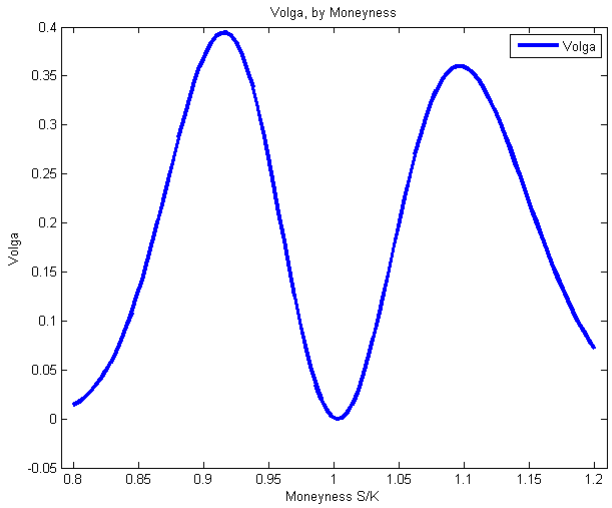



Figure 3.4 plots Black-Scholes volga (volatility gamma) by option moneyness for average levels of volatility.

# APPENDIX

## A.1. Appendix for Gold, Platinum, and Expected Stock Returns

### *A.1.1. Econometric Inference for Predictive Regressions*

Stambaugh (1999) shows that predictive regressions using persistent predictors are biased in finite samples. The standard return predictability regression is:

$$r^e_{t+1} = \alpha + \beta x_t + \epsilon_{t+1}$$

where $r^e_{t+1} = \log\left(\frac{P_{t+1}+D_{t+1}}{P_t}\right) - r^f_t$ is the log excess return from time $t$ to time $t+1$, and $x_t$ is some predictor known at time $t$ such as the log price-dividend ratio or the log GP ratio. If $x_t$ is a persistent predictor, we can model it as an AR(1) process:

$$x_{t+1} = \mu + \rho x_t + u_{t+1}$$

For predictors such as the price-dividend ratio, $cov(\epsilon, u) \neq 0$, since a positive return shock typically means prices increased, which also increases the price-dividend ratio. Letting $\gamma = \frac{cov(\epsilon_{t+1}, u_{t+1})}{var(u_{t+1})}$, the bias in the estimate of the predictive beta can be written as:

$$\mathbb{E}[\hat{\beta} - \beta] = \gamma \underbrace{\mathbb{E}[\hat{\rho} - \rho]}_{\approx -\frac{(1+3\rho)}{T} < 0} \tag{A.1}$$

The degree of bias is proportional to $\gamma$, which can be estimated as the slope of the regression of residuals from the predictive regression on the residuals from the AR(1) regression of the predictor variable. For the price-dividend ratio, the correlation between $\epsilon$ and $u$ in the data is 0.94, while it is only -0.17 for the GP ratio. Also note that there is no mechanical correlation between the residuals as is the case for the PD ratio. More formally, I project $\hat{\epsilon_t}$ on $\hat{u_t}$ and estimate $\hat{\gamma}$ to be 10.55 for the PD ratio, whereas for the GP ratio $\hat{\gamma}$ is only -1.78. Evaluating at the maximum bias ($\rho = 1$) estimates an upper bound of -0.090 for PD ratio

bias, which is enough to change the sign, whereas it is only 0.015 for the GP ratio, which is small compared to the predictive beta of 0.237. The evidence suggests that the GP ratio predictability is not driven by finite sample bias.

Predictor persistence also potentially affects the size of tests (see e.g., Torous et al., 2004). For $\delta = corr(\epsilon, u)$, the test statistic for $\beta$ has a non-standard limiting distribution:

$$t_\beta \Longrightarrow \delta\tau_\rho + \sqrt{(1-\delta^2)}\, z$$

where $\tau_\rho$ is non-normal and $z$ is normal. I follow Elliot and Stock (1994) and use Monte Carlo simulations to asses the magnitude of these size distortions. I run 100,000 simulations of length equal to my sample size at a monthly frequency by simulating the above dynamics, evaluating all parameters using their sample values. When $\delta = 0.94$ (which is the case for the PD ratio), a 5% test has a true rejection rate of 17%. For the GP ratio, where $\delta = -0.167$, a 5% test has a true rejection rate of 6%, which is very close. Since the absolute value of $\delta$ is small in the case of the GP ratio, the significance of the predictability tests is not affected by potential size distortions due to predictor persistence.

*A.1.2. Realized Utility Gains*

I calculate the realized utility gains for an investor who maximizes mean-variance preferences given a risk aversion of $\gamma$.[1] Given forecasts of expected returns and stock market volatility, the investor optimally allocates between stocks which earn the market return and bonds which earn the risk-free rate. The allocation to stocks for period $t+1$ is formed in period $t$. Using the historical average as the estimate of expected return, the allocation is:

$$w_{1,t} = \left(\frac{1}{\gamma}\right)\left(\frac{\overline{r}_{t+1}}{\widehat{\sigma}^2_{t+1}}\right)$$

[1]Following the literature, I set $\gamma = 3$.

Using the GP ratio, the allocation to stocks is given by:

$$w_{2,t} = \left(\frac{1}{\gamma}\right)\left(\frac{\widehat{r}_{t+1}}{\widehat{\sigma}^2_{t+1}}\right)$$

For both portfolio choice problems, $\widehat{\sigma}^2_{t+1}$ is the forecasted variance of stock returns over the next month, which I estimate following Li et al. (2013) by using a ten-year trailing window of monthly stock returns. The average utility levels for the investor over the out-of-sample period are given by:

$$\begin{aligned} U_1 &= \mu_1 - \frac{1}{2}\gamma\widehat{\sigma}^2_1 \\ U_2 &= \mu_2 - \frac{1}{2}\gamma\widehat{\sigma}^2_2 \end{aligned} \tag{A.2}$$

where $\mu_i$ is the sample mean of the return for the portfolio formed based on strategy $i$, where strategy 1 uses the historical average and strategy 2 uses the GP ratio. We can view the utility level as a certainty equivalent return for an investor with these preferences (Kandel and Stambaugh (1996)).[2] The utility gain of using the GP ratio over the historical average in percentage terms is given by $1200 \times (U_2 - U_1)$, which can be thought of as the management fee that an investor with mean-variance preferences would be willing to pay to access the GP ratio to generate return forecasts. In the data, the GP ratio produces a large, positive utility gain of 4.53% while other popular forecasting variables offer much lower or weakly negative utility gains.[3]

*A.1.3. Stationary Mean of $\lambda_t$*

I compute the stationary mean of the $\lambda_t$ proces following Nowotny (2011), adapted to my economic environment. The process is given by:

$$d\lambda_t = \kappa_\lambda(\xi_t - \lambda_t)dt + \sigma_\lambda\sqrt{\lambda_t}dW^\lambda_t + J^\lambda_t dN^\lambda_t$$

[2]Following Campbell and Thompson (2008), I constrain the allocation to stocks to be between 0% and 150%.

[3]The results for other forecasting variables are similar to the results in Li et al. (2013).

which implies that

$$\begin{aligned}\frac{\mathbb{E}[\lambda_t]}{dt} &= \kappa_\lambda(n(t) - m(t)) + \mu_\lambda\rho_0 + \mu_\lambda\rho_1 m(t) \\ \frac{\mathbb{E}[\xi_t]}{dt} &= \kappa_\xi\bar{\xi} - \kappa_\xi n(t)\end{aligned} \tag{A.3}$$

where $m(t) = \mathbb{E}[\lambda_t]$ and $n(t) = \mathbb{E}[\xi_t]$, with $n(t) \to \bar{\xi}$. Solving the ordinary differential equation for $m(t)$ implied by (A.3) results in the stationary mean of $\lambda_t$:

$$\mathbb{E}[\lambda_\infty] = \lim_{t\to\infty} m(t) = \frac{\kappa_\lambda\bar{\xi} + \mu_\lambda\rho_0}{\kappa_\lambda - \mu_\lambda\rho_1} \tag{A.4}$$

with necessary conditions $\kappa_\lambda > \mu_\lambda\rho_1$ and $\kappa_\lambda\bar{\xi} + \mu_\lambda\rho_0 > 0$, which are satisfied under the model calibration.

*A.1.4. Gold and Platinum Mine Production*

The data for world platinum mine production is from Johnson Matthey. The data for world gold mine production is from the U.S. Geological Survey (USGS) Annual Mineral Yearbook reports. The data is annual from 1975 to 2013. The Johnson Matthey data details both annual platinum mine production as well as autocatalyst demand and salvage. I use as my measure of the increment to the platinum stock the total quantity mined in a given year minus the autocatalyst demand net of salvage. Thomas and Boyle (1986) estimate the initial world above-ground stock of gold at the end of 1974 to be 84,000 tonnes (2,700 million troy oz). There is not a consensus estimate of world above-ground platinum stock (net of autocatalysts) that I am aware of, although during the 1975 to 2013 period in the data, annual platinum production (net of autocatalyst demand) is consistently approximately 4.5% of gold production with very little variation each year. Annual gold production is approximately 2,000 tonnes and platinum production is approximately 90 tonnes. Therefore, I estimate the initial stock of platinum to be 3,780 tonnes.[4] Using

[4]Is this a reasonable estimate? While the discovery of platinum is often credited to Antonio de Ulloa in 1735, it was not until Hans Merensky identified large economic deposits of platinum in the Bushveld Igneous Complex of South Africa in 1924 that large scale platinum mining took place (Cawthorn (1999)). My results are robust to reasonable perturbations of the initial estimate.

market prices at the end of 2013, this puts the total dollar value of all gold in the world at \$6.4 trillion, and the value of all platinum in the world (not found in autocatalysts) at just over \$300 billion. I proxy for population growth using U.S. annual population growth data provided by the U.S. Census Bureau.[5]

Panel A of Table 1.16 describes the log growth rate of the aggregate per-capita stock of gold and platinum. $G_t$ is the per-capita stock of gold, and $X_t$ is the per-capita stock of platinum. The mean per-capita log growth rate of the aggregate gold stock is 0.72% per year, and the growth rate is very smooth: the standard deviation is only 0.21%. The platinum stock displays similar dynamics, with an average growth rate of 0.71% and a standard deviation of 0.29%. Furthermore, the means are close to the medians.

Given the stable relationship between gold and platinum production each year, I look for evidence of cointegration between the log per-capita stock of gold and platinum. Two processes $\log X_t$ and $\log G_t$ are cointegrated if there exists a vector $\beta$ such that $\beta' \begin{bmatrix} \log X_t \\ \log G_t \end{bmatrix}$ is a stationary process. Panel B shows that $\log G_t$ and $\log X_t$ are unit root processes: an augmented Dickey and Fuller (1979) test fails to reject the null of a unit root at all lags 1 through 5. However, the process $\log Z_t = \log X_t - \log G_t$ appears to be stationary. I estimate the cointegration vector using Dynamic Least Squares (DLS) as suggested by Stock and Watson (1993)in Panel C of Table 1.16:

$$\log X_t = \beta_0 + \beta_G \log G_t + \sum_{i=-k}^{k} \gamma_i \Delta \log G_{t-i} + \epsilon_t \tag{A.5}$$

for $k = 1, 2, 3$. The estimates of $\beta_G$ are significant, ranging from 0.99 to 1.04, and in all cases a 95% confidence interval includes 1, which suggests that the cointegration vector is not statistically different from $[1, -1]$.

As further evidence of cointegration, I estimate the joint system $Y_t = \begin{bmatrix} \log X_t & \log G_t \end{bmatrix}'$ in

[5] I use U.S. population growth as opposed to world population growth to be consistent with the consumption data in the calibration, which uses U.S. per-capita consumption data.

a Engle and Granger (1987) Vector Error-Correction Model (VECM):

$$\Delta Y_t = \mu + \Pi Y_{t-1} + \sum_{j=1}^{p-1} \Gamma_j \Delta Y_{t-j} + \epsilon_t \tag{A.6}$$

and conduct Johansen (1988) rank tests for cointegration based on the rank of the matrix $\Pi$. The null hypothesis for the rank test is that there are no more than $r$ cointegrating relationships, which implies that the remaining $K - r$ eigenvalues of $\Pi$ must be zero where $K$ is the dimension of $Y_t$. I follow Johansen (1995) and apply an iterative procedure which starts testing at $r = 0$ and accepts as $\hat{r}$ (number of cointegrating relationships) the first value of $r$ for which the test fails to reject the null. Table 1.16 Panel D shows the results for the VECM with 1 through 4 lags. We see that for the estimation with 3 lags (which is the optimal lag length as chosen by the Akaike Information Criterion), we reject the null of zero cointegrating relationships, but fail to reject the null of 1 cointegrating relationship.[6] Figure 1.9 plots the demeaned $\log X_t$ and $\log G_t$ processes, where we can clearly see that gold and platinum supply seem to track each other over time.

*A.1.5. Model Solution*

The equations for the Epstein-Zin discount factor coefficients are given by:

$$\begin{aligned}
a &= \frac{(1-\gamma)(\frac{1}{2}(1-\gamma)\sigma_c^2)}{\delta} + \frac{b_\xi \kappa_\xi \bar{\xi} + \rho_0 \mathbb{E}_\eta \left[ e^{b_\lambda J_t^\lambda} - 1 \right]}{\delta} \\
0 &= \frac{1}{2}\sigma_\lambda^2 {b_\lambda}^2 - (\kappa_\lambda + \delta) b_\lambda + \rho_1 \mathbb{E}_\eta \left[ e^{b_\lambda J_t^\lambda} - 1 \right] + \mathbb{E}_v \left[ e^{(1-\gamma) J_t^c} - 1 \right] \\
b_\xi &= \frac{k_\xi + \delta}{{\sigma_\xi}^2} - \sqrt{\left( \frac{k_\xi + \delta}{{\sigma_\xi}^2} \right)^2 - \frac{2 b_\lambda \kappa_\lambda}{{\sigma_\xi}^2}}.
\end{aligned}$$

In general, I can allow $\lambda_t^\lambda = \rho_0 + \rho_1 \lambda_t$. Equity price-dividend ratio is given by:

$$\frac{P_t}{D_t} = \int_0^\infty \exp(a_\phi(\tau) + b_\phi(\tau)\lambda_t + c_\phi(\tau)\xi_t) d\tau$$

[6] The estimated $\beta_G$ in the VECM from $\Pi = \alpha\beta'$ has a 95% confidence interval of [-1.45,-1.04].

with $a_\phi(\tau), b_\phi(\tau), c_\phi(\tau)$ given by the ODEs:

$$\begin{aligned}
a'_\phi(\tau) &= c_\phi(\tau)\kappa_\xi\bar{\xi} + \phi\bar{g}_c + \frac{1}{2}\phi^2\sigma_c^2 - \delta - (\bar{g}_c + \frac{1}{2}\sigma_c^2) \\
&\quad + \gamma(1-\phi)\sigma_c^2 + \rho_0\mathbb{E}_\eta\left[e^{(b_\lambda + b_\phi(\tau))J_t^\lambda} - e^{b_\lambda J_t^\lambda}\right] \\
b'_\phi(\tau) &= \frac{1}{2}\sigma_\lambda^2 b_\phi(\tau)^2 + (b_\lambda\sigma_\lambda^2 - \kappa_\lambda)b_\phi(\tau) + \mathbb{E}_v\left[e^{(\phi-\gamma)J_t^c} - e^{(1-\gamma)J_t^c}\right] \\
&\quad + \rho_1\mathbb{E}_\eta\left[e^{(b_\lambda + b_\phi(\tau))\mathbb{J}_t^\lambda} - e^{b_\lambda J_t^\lambda}\right] \\
c'_\phi(\tau) &= \frac{1}{2}\sigma_\xi^2 c_\phi(\tau)^2 + (b_\xi\sigma_\xi^2 - \kappa_\xi)c_\phi(\tau) + b_\phi(\tau)\kappa_\lambda
\end{aligned}$$

with initial conditions $a_\phi(0) = b_\phi(0) = c_\phi(0) = 0$.

Let $P_{g,t}^{t+\tau}$ be the price of zero-coupon gold which pays $Q_{g,t+\tau}$ and nothing else, and let $P_{x,t}^{t+\tau}$ be the analogous claim for platinum. Gold and platinum price-dividend ratios are solved by noting that $\pi_t P_{g,t}^{t+\tau}$ and $\pi_t P_{x,t}^{t+\tau}$ are martingales, so the sum of the drift and jump compensator must equal zero.

The gold price-dividend ratio is given by:

$$\frac{P_{g,t}}{Q_{g,t}} = \int_0^\infty \exp(a_g(\tau) + b_g(\tau)\lambda_t + c_g(\tau)\xi_t)d\tau$$

with $a_g(\tau), b_g(\tau), c_g(\tau)$ given by the ODEs:

$$\begin{aligned}
a'_g(\tau) &= c_g(\tau)\kappa_\xi\bar{\xi} + \frac{1}{\epsilon}\left[\bar{g}_c - \mu_g + \frac{1}{2\epsilon}(\sigma_c^2 + \sigma_g^2)\right] - \delta - (\bar{g}_c + \frac{1}{2}\sigma_c^2) \\
&\quad + \gamma(1 - \frac{1}{\epsilon})\sigma_c^2 + \rho_0\mathbb{E}_\eta\left[e^{(a_2 + b_\lambda + b_g(\tau))J_t^\lambda} - e^{b_\lambda J_t^\lambda}\right] \\
b'_g(\tau) &= \frac{1}{2}\sigma_\lambda^2 b_g(\tau)^2 + \left[(a_2 + b_\lambda)\sigma_\lambda^2 - \kappa_\lambda\right]b_g(\tau) \\
&\quad + \frac{1}{2}\sigma_\lambda^2 a_2^2 + a_2(b_\lambda\sigma_\lambda^2 - \kappa_\lambda) + \mathbb{E}_v\left[e^{(\frac{1}{\epsilon}-\gamma)J_t^c} - e^{(1-\gamma)J_t^c}\right] \\
&\quad + \rho_1\mathbb{E}_\eta\left[e^{(a_2 + b_\lambda + b_g(\tau))\mathbb{J}_t^\lambda} - e^{b_\lambda J_t^\lambda}\right] \\
c'_g(\tau) &= \frac{1}{2}\sigma_\xi^2 c_g(\tau)^2 + (b_\xi\sigma_\xi^2 - \kappa_\xi)c_g(\tau) + (a_2 + b_g(\tau))\kappa_\lambda
\end{aligned}$$

with initial conditions $a_g(0) = b_g(0) = c_g(0) = 0$.

The platinum price-dividend ratio is given by:

$$\frac{P_{x,t}}{Q_{x,t}} = \int_0^\infty \exp(a_x(\tau) + b_x(\tau)\lambda_t + c_x(\tau)\xi_t + d_x(\tau)\log Z_t)d\tau$$

with $a_x(\tau), b_x(\tau), c_x(\tau), d_x(\tau)$ given by the ODEs:

$$\begin{aligned}
a_x'(\tau) &= c_x(\tau)\kappa_\xi\bar{\xi} + \frac{1}{\epsilon}\left[\bar{g_c} - \mu_g + \frac{1}{2\epsilon}(\sigma_c^2 + \sigma_g^2 + \sigma_z^2)\right] - \delta - (\bar{g_c} + \frac{1}{2}\sigma_c^2) \\
&\quad + \gamma(1 - \frac{1}{\epsilon})\sigma_c^2 + \frac{1}{2}\sigma_z^2 d_x(\tau)^2 + d_x(\tau)(\theta_z\mu_z - \frac{1}{\epsilon}\sigma_z^2) \\
&\quad - \frac{1}{\epsilon}\theta_z\mu_z + \rho_0\mathbb{E}_\eta\left[e^{(b_2+b_\lambda+b_x(\tau))J_t^\lambda} - e^{b_\lambda J_t^\lambda}\right] \\
b_x'(\tau) &= \frac{1}{2}\sigma_\lambda^2 b_x(\tau)^2 + \left[(b_2 + b_\lambda)\sigma_\lambda^2 - \kappa_\lambda\right] b_x(\tau) \\
&\quad + \frac{1}{2}\sigma_\lambda^2 b_2^2 + b_2(b_\lambda\sigma_\lambda^2 - \kappa_\lambda) + \mathbb{E}_v\left[e^{(\frac{1}{\epsilon}-\gamma)J_t^c} - e^{(1-\gamma)J_t^c}\right] \\
&\quad + \rho_1\mathbb{E}_\eta\left[e^{(b_2+b_\lambda+b_x(\tau))\mathrm{J}_t^\lambda} - e^{b_\lambda J_t^\lambda}\right] \\
c_x'(\tau) &= \frac{1}{2}\sigma_\xi^2 c_x(\tau)^2 + (b_\xi\sigma_\xi^2 - \kappa_\xi)c_x(\tau) + (b_2 + b_x(\tau))\kappa_\lambda \\
d_x'(\tau) &= -\theta_z d_x(\tau) + \theta_z\frac{1}{\epsilon}
\end{aligned}$$

with initial conditions $a_x(0) = b_x(0) = c_x(0) = d_x(0) = 0$.

*A.1.6. Log-Linearized Gold and Platinum Prices*

Gold price-dividend ratios are given by

$$\frac{P_{g,t}}{Q_{g,t}} = G^g(\lambda_t, \xi_t) = \int_0^\infty e^{a_g(\tau)+b_g(\tau)\lambda_t+c_g(\tau)\xi_t}d\tau$$

Let $g(\lambda_t, \xi_t) = \log G^g(\lambda_t, \xi_t)$. Given fixed $\lambda^*, \xi^*$, Taylor expansion implies that

$$g(\lambda, \xi) \approx g(\lambda^*, \xi^*) + \left.\frac{\partial g}{\partial\lambda}\right|_{(\lambda^*,\xi^*)}(\lambda_t - \lambda^*) + \left.\frac{\partial g}{\partial\xi}\right|_{(\lambda^*,\xi^*)}(\xi_t - \xi^*)$$

where we have that

$$\left.\frac{\partial g}{\partial \lambda}\right|_{(\lambda^*,\xi^*)} = \frac{1}{G(\lambda^*,\xi^*)}\int_0^\infty b_g(\tau)e^{a_g(\tau)+b_g(\tau)\lambda_t+c_g(\tau)\xi_t}d\tau = b^*_{g,\lambda}$$
$$\left.\frac{\partial g}{\partial \xi}\right|_{(\lambda^*,\xi^*)} = \frac{1}{G(\lambda^*,\xi^*)}\int_0^\infty c_g(\tau)e^{a_g(\tau)+b_g(\tau)\lambda_t+c_g(\tau)\xi_t}d\tau = b^*_{g,\xi}.$$

This implies that $G^g(\lambda_t,\xi_t) \approx G^g(\lambda^*,\xi^*)e^{b^*_{g,\lambda}(\lambda_t-\lambda^*)+b^*_{g,\xi}(\xi_t-\xi^*)}$, and I set $\lambda^*$ and $\xi^*$ equal to the stationary means of $\lambda_t$ and $\xi_t$, respectively. Since $Q_{g,t} = e^{a_1+a_2\lambda_t}e^{\frac{1}{\epsilon}\log C_t-\frac{1}{\epsilon}\log G_t}$, This implies that log-linearized gold prices are given by

$$\log P_{g,t} = A_g + \frac{1}{\epsilon}\log C_t - \frac{1}{\epsilon}\log G_t + (a_2 + b^*_{g,\lambda})\lambda_t + b^*_{g,\xi}\xi_t.$$

Similarly, log-linearized platinum prices are given by

$$\log P_{x,t} = A_x + \frac{1}{\epsilon}\log C_t - \frac{1}{\epsilon}\log G_t + (b_2 + b^*_{x,\lambda})\lambda_t + b^*_{x,\xi}\xi_t + (b^*_{x,Z} - \frac{1}{\epsilon})\log Z_t.$$

The constants $A_g$ and $A_x$ only affect the level of GP and are mainly determined by the scaling term $a_1 - b_1$.

*A.1.7. Implied Volatilities in the Model*

While the model implies that the GP ratio captures time-variation in the tail risk measures $\lambda_t$ and $\xi_t$, can it formally reconcile the empirical evidence that the GP ratio can explain the slope of the implied volatility curve for equity index options? To investigate this, I price options in the model following Eraker and Shaliastovich (2008), Nowotny (2011), and Seo and Wachter (2014). Since the model is in the class of affine jump diffusion models studied by Duffie et al. (2000), the solution for the discounted characteristic function is known up to a system of differential equations.The risk-neutral dynamics of the state variables are also in the class of affine jump diffusions and we can derive the discounted characteristic

function where $X_t$ is a vector of the state variables:

$$\mathbb{E}_t^{\mathbb{Q}}\left[e^{-\int_t^{t+\tau} r_s^f ds} e^{uX_{t+\tau}}\right] = e^{\alpha(\tau)+\beta(\tau)X_t} \tag{A.7}$$

and $\alpha(\tau), \beta(\tau)$ are the solutions to a system of ordinary differential equations as given in Duffie et al. (2000). Nowotny (2011) describes the change of measure in detail.

I approximate the price-dividend ratio using a Taylor approximation following Seo and Wachter (2014), who also show that the approximation is very accurate when the intertemporal elasticity of substitution is equal to one. This allows me to express the price-dividend ratio as:

$$G(\lambda_t, \xi_t) \approx G(\lambda^*, \xi^*) e^{b^*_{\phi,\lambda}(\lambda_t-\lambda^*)+b^*_{\phi,\xi}(\xi_t-\xi^*)} \tag{A.8}$$

for constants $\lambda^*$ and $\xi^*$ which I set equal to the long-run mean of $\lambda_t$ and $\xi_t$, respectively. Let $A_0 = \log G(\lambda^*, \xi^*) - b^*_{\phi,\lambda}\lambda^* - b^*_{\phi,\xi}\xi^*$. We can express the price of equity at time $t+\tau$ as:

$$P_{t+\tau} = e^{A_0 + d' X_{t+\tau}}, \text{ for } d = [\phi, b^*_{\phi,\lambda}, b^*_{\phi,\xi}]'$$

Following Carr and Madan (1999), Lewis (2000), and Eraker and Shaliastovich (2008), I price options using Fourier transform methods. The forward and inverse Fourier transforms are:

$$\hat{f}(z) = \int_{-\infty}^{\infty} f(x) e^{izx} dx$$
$$f(x) = \frac{1}{2\pi} \int_{iz_i-\infty}^{iz_i+\infty} \hat{f}(z) e^{-izx} dz$$

where $i^2 = -1$ and $z_i$ denotes the imaginary part of the complex variable $z$. The payoff of a European put option is given by $f(x) = \max(K - e^x, 0)$, whose inverse transform is given

by:

$$\hat{f}(z) = -K\frac{K^{iz}}{z^2 - iz}$$

The price of a European put option is a function of the state variables $X_t$, strike price $K$, and time to maturity $\tau$, and can be computed as:

$$\begin{aligned} P(X_t, K, \tau) &= \mathbb{E}_t^{\mathbb{Q}}\left[e^{-\int_t^{t+\tau} r_s^f ds}\max(K - e^{A_0 + d'X_{t+\tau}}, 0)\right] \\ &= \frac{-K}{2\pi}\int_{iz_i-\infty}^{iz_i+\infty} e^{izA_0}\mathbb{E}_t^{\mathbb{Q}}\left[e^{-\int_t^{t+\tau} r_s^f ds}e^{-izdX_{t+\tau}}\right]\left(\frac{K^{iz}}{z^2 - iz}\right)dz \\ &= \frac{-K}{2\pi}\int_{iz_i-\infty}^{iz_i+\infty} e^{izA_0}e^{\alpha(\tau)+\beta(\tau)X_t}\left(\frac{K^{iz}}{z^2 - iz}\right)dz \end{aligned} \tag{A.9}$$

with $\alpha(0) = 0$, $\beta(0) = -izd$, and $z_i < 0$. I normalize strike price and option price by $P_t$ so strike prices can be interpreted as moneyness. Given put option prices, I back out Black and Scholes (1973b) implied volatilities using the endogenous dividend yield $\frac{1}{G(\lambda_t,\xi_t)}$ and set the risk-free rate equal to the government bond rate in the model.

Figure 1.10 plots the difference between the implied volatility of a OTM put option with 0.95 moneyness and an ATM put option with moneyness equal to 1, as a function of the state variables $\lambda_t$ and $\xi_t$. The options have 1 month to maturity as in the data. This figure is quite informative because it gives intuition about how $\lambda_t$ and $\xi_t$ affect return volatility (endogenously determined) and tail risk in the model.[7] In a one-factor model, where $\lambda_t$ reverts to a constant, both volatility and disaster risk are controlled by the same variable. The combined result is that when $\lambda_t$ increases, this increases the average level of volatility in addition to the likelihood of negative jumps. Option maturities are typically short (measured in months, for most liquid equity index options), and at this horizon the increase in volatility level is actually stronger than the increase in tail risk, resulting in the implied volatility slope being a decreasing function of $\lambda_t$ in the one-factor model.

[7] Seo and Wachter (2014) offer a detailed discussion of this topic.

In the two-factor model, $\xi_t$ controls the level of volatility more than it affects the likelihood of jumps, which is directly controlled by $\lambda_t$. Therefore, we see that the implied volatility slope is increasing in $\lambda_t$, and particularly fast when $\xi_t$ is low. Similarly, the implied volatility slope decreases in $\xi_t$, and especially quickly when $\lambda_t$ and $\xi_t$ are low. As shown earlier, GP loads positively on both $\lambda_t$ and $\xi_t$. For computational tractability, I first compute the implied volatility slope over a fine mesh of the state variables $(\lambda_t, \xi_t)$. I then simulate repeated samples of length equal to the data counterpart (17 years). The model implied volatility slope is calculated by interpolating the mesh surface, and regressed on GP in the model. We see a coefficient as large as the data estimate in over 20% of sample paths. Staying within the class of rare disaster models, the two-factor model is necessary to allow the disaster intensity and the implied volatility slope to be positively related.

## A.2. Appendix for Risk Adjustment and the Temporal Resolution of Uncertainty

### *A.2.1. Model Solution*

The equilibrium price of the asset is computed using the Euler equation:

$$\mathbb{E}_i \left[ M_{i,j} R_j \right] = 1. \tag{A.10}$$

We first use this equation for the consumption asset. Given the expression for the stochastic discount factor in (2.12), we obtain that in equilibrium,

$$E_i \left[ \delta^\theta \lambda_j^{1-\gamma} \left( \frac{PC_j + 1}{PC_i} \right)^\theta \right] = 1. \tag{A.11}$$

This provides us the equation for the wealth-to-consumption ratio in each state $i$, and we solve the system of equations numerically using fixed point iteration.

Given equilibrium solutions to the wealth-to-consumption ratio, we can characterize the stochastic discount factor and obtain equilibrium prices of the Arrow-Debreu claims, the stock market, and the risk-free rates. Specifically, the Arrow-Debreu prices follow the equation (2.1). For the stock market claim, the Euler equation is given by,

$$E_i \left[ M_{i,j} \left( PD_j + 1 \right) \left( \mu_d + \phi(\lambda_j - \mu) \right) \right] = PD_i, \tag{A.12}$$

which leads to a linear matrix system for the price-dividend ratio of the market, $PD_i$. Risk-free rates satisfies

$$R_{f,i} = E_i \left[ M_{i,j} \right]^{-1}, \tag{A.13}$$

which can be solved using the equilibrium dynamics for the stochastic discount factor $M_{i,j}$.

*A.2.2. Estimation of Risk-Neutral Density*

The exact steps for the estimation of the risk-neutral density from options prices are provided below:

1. Combine puts and calls, which are out-of-the-money (not too deep out-of-the-money, best bid at least $0.50), and contracts not more than 20 points in-the-money.

2. Transform mid-prices into implied volatilities using Black and Scholes (1973a). In the region of +/- 20 points from at-the-money, take a weighted average of put and call implied volatilities.

3. Fit a 4th order polynomial to the implied volatilities over a dense set of strike prices, and convert back into call option prices using Black-Scholes.

4. Numerically differentiate the call prices using (3.3) to recover the risk-neutral distribution function.

## A.3. Appendix for Volatility-of-Volatility Risk

### *A.3.1. Delta-Hedged Equity Options*

The state vector is $x_t = \begin{bmatrix} S_t & V_t & \eta_t \end{bmatrix}'$. Under the linear risk premium structure, $\lambda_t^V = \lambda^V V_t$ and $\lambda_t^\eta = \lambda^\eta \eta_t$. Note that since $C_t$ is homogeneous of degree 1 in the underlying $S_t$, $\frac{\partial C}{\partial V}$ and $\frac{\partial C}{\partial \eta}$ are also homogeneous of degree 1 in $S_t$. Define a pair of functions:

$$\begin{aligned} g_1(x_t) &= \lambda_t^V \frac{\partial C_t}{\partial V_t} = \lambda^V V_t h_t^1(\tau; y) S_t \\ g_2(x_t) &= \lambda_t^\eta \frac{\partial C_t}{\partial \eta_t} = \lambda^\eta \eta_t h_t^2(\tau; y) S_t. \end{aligned} \tag{A.14}$$

We can re-write equation (3.8) as:

$$\mathbb{E}_t \left[ \Pi_{t,t+\tau} \right] = \mathbb{E}_t \left[ \int_t^{t+\tau} g_1(x_u) du \right] + \mathbb{E}_t \left[ \int_t^{t+\tau} g_2(x_u) du \right]. \tag{A.15}$$

Define operators $\mathcal{L}$ and $\Gamma$ such that:

$$\begin{aligned} \mathcal{L}[.]\, dt = & \frac{\partial [.]}{\partial S} \mu_t S_t dt + \frac{\partial [.]}{\partial V} \theta(V_t) dt + \frac{\partial [.]}{\partial \eta} \gamma(\eta_t) dt + \frac{\partial [.]}{\partial t} dt \\ & + \frac{1}{2} \frac{\partial^2 [.]}{\partial S^2} [dS_t, dS_t] + \frac{1}{2} \frac{\partial^2 [.]}{\partial V^2} [dV_t, dV_t] + \frac{1}{2} \frac{\partial^2 [.]}{\partial \eta^2} [d\eta_t, d\eta_t] \\ & + \frac{\partial^2 [.]}{\partial S \partial \eta} [dS_t, d\eta_t] + \frac{\partial^2 [.]}{\partial S \partial V} [dS_t, dV_t] + \frac{\partial^2 [.]}{\partial V \partial \eta} [dV_t, d\eta_t] \\ \Gamma[.] = & \left[ \frac{\partial [.]}{\partial S} S_t \sqrt{V_t},\ \frac{\partial [.]}{\partial V} \sqrt{\eta_t},\ \frac{\partial [.]}{\partial \eta} \phi \sqrt{\eta_t} \right]. \end{aligned} \tag{A.16}$$

Then, for $u > t$, Itô's Lemma implies that:

$$g_1(x_u) = g_1(x_t) + \int_t^u \mathcal{L} g(x_{u'}) du' + \int_t^u \Gamma g(x_{u'}) dW_{u'}.$$

The integral in the first expectation on the right-hand side of equation (A.15) becomes:

$$\begin{aligned}\int_t^{t+\tau} g_1(x_u)du &= \int_t^{t+\tau}\left[g_1(x_t) + \int_t^u \mathcal{L}g(x_{u'})du' + \int_t^u \Gamma g(x_{u'})dW_{u'}\right]du \\ &= g_1(x_t)\tau + \frac{1}{2}\mathcal{L}g_1(x_t)\tau^2 + \frac{1}{6}\mathcal{L}^2 g_1(x_t)\tau^3 + ... + \text{Itô Integrals} \\ &= \sum_{n=0}^{\infty}\frac{\tau^{1+n}}{(1+n)!}\mathcal{L}^n g_1(x_t) + \text{Itô Integrals},\end{aligned}$$

and likewise for the second integral in (A.15). We can use this to re-write (A.15) as:

$$\begin{aligned}\mathbb{E}_t\left[\Pi_{t,t+\tau}\right] &= \mathbb{E}_t\left[\int_t^{t+\tau} g_1(x_u)du\right] + \mathbb{E}_t\left[\int_t^{t+\tau} g_2(x_u)du\right] \\ &= \sum_{n=0}^{\infty}\frac{\tau^{1+n}}{(1+n)!}\mathcal{L}^n\left[g_1(x_t)\right] + \sum_{n=0}^{\infty}\frac{\tau^{1+n}}{(1+n)!}\mathcal{L}^n\left[g_2(x_t)\right].\end{aligned} \tag{A.17}$$

Note that $g_1(x_t) = \alpha_1(V_t, \tau; y)S_t$, and $g_2(x_t) = \alpha_2(\eta_t, \tau; y)S_t$. By Lemma 1 of Bakshi and Kapadia (2003), $\mathcal{L}^n[g_1(x_t)]$ and $\mathcal{L}^n[g_2(x_t)]$ will also be proportional to $S_t$, which implies that:

$$\begin{aligned}\mathcal{L}^n\left[g_1(x_t)\right] &= \lambda^V V_t \Phi^V_{t,n} S_t \ \ \forall n \\ \mathcal{L}^n\left[g_2(x_t)\right] &= \lambda^\eta \eta_t \Phi^\eta_{t,n} S_t \ \ \forall n.\end{aligned}$$

Therefore, we have:

$$\begin{aligned}\mathbb{E}_t\left[\Pi_{t,t+\tau}\right] &= \sum_{n=0}^{\infty}\frac{\tau^{1+n}}{(1+n)!}\mathcal{L}^n\left[g_1(x_t)\right] + \sum_{n=0}^{\infty}\frac{\tau^{1+n}}{(1+n)!}\mathcal{L}^n\left[g_2(x_t)\right] \\ &= S_t\left[\lambda^V\beta_t^V V_t + \lambda^\eta\beta_t^\eta\eta_t\right],\end{aligned}$$

which implies that:

$$\frac{\mathbb{E}_t\left[\Pi_{t,t+\tau}\right]}{S_t} = \lambda^V\beta_t^V V_t + \lambda^\eta\beta_t^\eta\eta_t, \tag{A.18}$$

where the sensitivities to the risk factors are given by:

$$\begin{aligned}\beta_t^V &= \sum_{n=0}^{\infty} \frac{\tau^{1+n}}{(1+n)!}\Phi_{t,n}^V > 0 \\ \beta_t^\eta &= \sum_{n=0}^{\infty} \frac{\tau^{1+n}}{(1+n)!}\Phi_{t,n}^\eta > 0.\end{aligned} \tag{A.19}$$

The betas are positive since $\frac{\partial C_t}{\partial V_t} > 0$ and $\frac{\partial C_t}{\partial \eta_t} > 0$.

*A.3.2. Delta-Hedged VIX Options*

The state vector is $x_t = \begin{bmatrix} F_t & \eta_t \end{bmatrix}'$, and $g(x_t) = \frac{\partial C_t^*}{\partial \eta_t}\lambda_t^\eta$. Again, we will apply Itô-Taylor expansions. Let operators $\mathcal{L}$ and $\Gamma$ be such that:

$$\begin{aligned}\mathcal{L}[.]\, dt &= \frac{\partial[.]}{\partial F}\mu_{F,t}F_t dt + \frac{\partial[.]}{\partial \eta}\gamma(\eta_t)dt + \frac{\partial[.]}{\partial t}dt \\ &\quad + \frac{1}{2}\frac{\partial^2[.]}{\partial F^2}[dF_t, dF_t] + \frac{1}{2}\frac{\partial^2[.]}{\partial \eta^2}[d\eta_t, d\eta_t] + \frac{\partial^2[.]}{\partial F \partial \eta}[dF_t, d\eta_t] \\ \Gamma[.] &= \begin{bmatrix}\frac{\partial[.]}{\partial F}\sigma_{F_t},\ \frac{\partial[.]}{\partial \eta}\phi\sqrt{\eta_t}\end{bmatrix}.\end{aligned} \tag{A.20}$$

By Itô's Lemma, we have:

$$g(x_u) = g(x_t) + \int_t^u \mathcal{L}g(x_{u'})du' + \int_t^u \Gamma g(x_{u'})dW_{u'}.$$

Then, we have that:

$$\begin{aligned}\int_t^{t+\tau} g(x_u)du &= \int_t^{t+\tau}\left(g(x_t) + \int_t^u \mathcal{L}g(x_{u'})du' + \int_t^u \Gamma g(x_{u'})dW_{u'}\right)du \\ &= \int_t^{t+\tau} g(x_t)du + \int_t^{t+\tau}\int_t^u \mathcal{L}g(x_{u'})du'du + \int_t^{t+\tau}\int_t^u \Gamma g(x_{u'})dW_{u'}du \\ &= g(x_t)\tau + \frac{1}{2}\mathcal{L}g(x_t)\tau^2 + \frac{1}{6}\mathcal{L}^\in g(x_t)\tau^3 + ... + \text{Itô integrals} \\ &= \sum_{n=0}^{\infty}\frac{\tau^{1+n}}{(1+n)!}\mathcal{L}^n g(x_t) + \text{Itô integrals}.\end{aligned} \tag{A.21}$$

This implies that $\mathbb{E}_t\left[\Pi_{t,t+\tau}\right] = \sum_{n=0}^{\infty} \frac{\tau^{1+n}}{(1+n)!}\mathcal{L}^n g(x_t)$, since the expectation of Itô integrals is zero.

$$\mathbb{E}_t\left[\Pi^*_{t,t+\tau}\right] = \mathbb{E}_t\left[\int_t^{t+\tau} \frac{\partial C^*_s}{\partial \eta_s}\lambda^\eta_s ds\right].$$

Re-write equation (3.18) using:

$$\begin{aligned} g(x_t) &= \lambda^\eta_t \frac{\partial C^*_t}{\partial \eta_t} \\ &= \lambda^\eta \eta_t h^2_t(\tau; y) F_t, \end{aligned}$$

where the second line follows from the homogeneity of $\frac{\partial C^*_t}{\partial \eta}$ in $F_t$, so $g(x_t) = \alpha(\eta_t, \tau; y)F_t$. By Lemma 1 of Bakshi and Kapadia (2003), $\mathcal{L}^n[g(x_t)]$ is also proportional to $F_t$, and we have that:

$$\begin{aligned} \mathbb{E}_t\left[\Pi_{t,t+\tau}\right] &= \mathbb{E}_t\left[\int_t^{t+\tau} \frac{\partial C_s}{\partial \eta_s}\lambda^\eta_s ds\right] \\ &= \mathbb{E}_t\left[\int_t^{t+\tau} g(x_s) ds\right] \\ &= \lambda^\eta \eta_t F_t \beta^*_t, \end{aligned} \tag{A.22}$$

where $\beta^*_t = \sum_{n=0}^{\infty} \frac{\tau^{1+n}}{(1+n)!}\Phi_{n,t}$ which is positive since $\frac{\partial C^*_t}{\partial \eta} > 0$. This gives us the familiar factor model structure:

$$\frac{\mathbb{E}_t\left[\Pi_{t,t+\tau}\right]}{F_t} = \lambda^\eta \beta^*_t \eta_t. \tag{A.23}$$

*A.3.3. Risk-Neutral Skewness*

The prices of the volatility, cubic, and quartic contracts $V_{t,t+\tau}, W_{t,t+\tau}, X_{t,t+\tau}$ are given

$$
\begin{aligned}
V_{t,t+\tau} &= \int_{S_t}^{\infty} \frac{2(1-\log\frac{K}{S_t})}{K^2} C(t,t+\tau;K)dK + \int_0^{S_t} \frac{2(1+\log\frac{S_t}{K})}{K^2} P(t,t+\tau;K)dK, \\
W_{t,t+\tau} &= \int_{S_t}^{\infty} \frac{6\log\frac{K}{S_t} - 3(\log\frac{K}{S_t})^2}{K^2} C(t,t+\tau;K)dK \\
&\quad - \int_0^{S_t} \frac{6\log\frac{S_t}{K} + 3(\log\frac{S_t}{K})^2}{K^2} P(t,t+\tau;K)dK, \\
X_{t,t+\tau} &= \int_{S_t}^{\infty} \frac{12(\log\frac{K}{S_t})^2 - 4(\log\frac{K}{S_t})^3}{K^2} C(t,t+\tau;K)dK \\
&\quad + \int_0^{S_t} \frac{12(\log\frac{S_t}{K})^2 + 4(\log\frac{S_t}{K})^3}{K^2} P(t,t+\tau;K)dK,
\end{aligned}
$$

and $\mu_{t,t+\tau} = e^{r_f\tau} - 1 - \frac{e^{r_f\tau}}{2} V_{t,t+\tau} - \frac{e^{r_f\tau}}{6} W_{t,t+\tau} - \frac{e^{r_f\tau}}{24} X_{t,t+\tau}$.

To construct these measures, we use out-of-the-money options to mitigate liquidity concerns. Following Shimko (1993), each day we interpolate the Black-Scholes implied volatility curve at the observable strikes using a cubic spline, and then calculate option prices to compute the above moments. We construct these measures for both S&P500 options and VIX options. Our implied volatility slope and risk-neutral skewness measures are calculated using options with the same maturity as our test assets.

# BIBLIOGRAPHY


A. Abel. Risk premia and term premia in general equilibrium. *Journal of Monetary Economics* 43, 3–33, 1999.

Y. Aït-Sahalia and A. W. Lo. Nonparametric risk management and implied risk aversion. *Journal of Econometrics* 94, 9–51, 2000.

Y. Aït-Sahalia, J. Parker, and M. Yogo. Luxury goods and the equity premium. *Journal of Finance* 59, 2959–3004, 2004.

S. Alizadeh, M. Brandt, and F. X. Diebold. Range-based estimation of stochastic volatility models. *Journal of Finance* 57, 1047–1091, 2002.

T. Andersen, N. Fusari, and V. Todorov. Parametric inference and dynamic state recovery from option panels. Working paper, 2012.

A. Ang and G. Bekaert. Stock Return Predictability: Is It There? *Review of Financial Studies* 20, 651–707, 2007.

O. Attanasio and G. Weber. Intertemporal substitution, risk aversion and the euler equation for consumption. *Economic Journal* 99, 59–73, 1989.

D. Backus, M. Chernov, and I. Martin. Disasters implied by equity index options. *Journal of Finance* 66, 1969–2012, 2011.

G. Bakshi and N. Kapadia. Delta-hedged gains and the negative market volatility risk premium. *Review of Financial Studies* 16, 527–566, 2003.

G. Bakshi, N. Kapadia, and D. Madan. Stock return characteristics, skew laws, and the differential pricing of individual equity options. *Review of Financial Studies* 16, 101–143, 2003.

R. Bansal, R. Dittmar, and C. Lundblad. Consumption, dividends, and the cross-section of equity returns. *Journal of Finance* 60, 1639–1672, 2005.

R. Bansal, D. Kiku, I. Shaliastovich, and A. Yaron. Volatility, the macroeconomy and asset prices. *Journal of Finance* 69, 2013.

R. Bansal, D. Kiku, and A. Yaron. Risks for the long run: Estimation and inference. Working paper, 2007.

R. Bansal, D. Kiku, and A. Yaron. An empirical evaluation of the long-run risks model for asset prices. *Critical Finance Review* 1, 183–221, 2011.

R. Bansal and I. Shaliastovich. Confidence risk and asset prices. *American Economic Review: Papers and Proceedings* 100, 537–541, 2011.

R. Bansal and A. Yaron. Risks for the long run: A potential resolution of asset pricing puzzles. *Journal of Finance* 54, 1481–1509, 2004.

O. Barndorff-Nielsen and N. Shephard. Power and bipower variation with stochastic volatility and jumps. *Journal of Financial Econometrics* 2, 1–37, 2004.

O. Barndorff-Nielsen and N. Shephard. Econometrics of testing for jumps in financial economics using bipower variation. *Journal of Financial Econometrics* 4, 1–30, 2006.

R. Barro. Rare disasters and asset markets in the twentieth century. *Quarterly Journal of Economics* 121, 823–866, 2006.

R. Barro and S. Misra. Gold returns. Working paper, 2013.

R. Barro and J. Ursua. Macroeconomic crises since 1870. *Brookings Papers on Economic Activity* 255–350, 2008.

Basel I. International convergence of capital measurement and capital standards. Basel Committee on Banking Supervision, Bank of International Settlements , 1988.

Basel II. Basel II: International convergence of capital measurement and capital standards: a revised framework. Basel Committee on Banking Supervision, Bank of International Settlements , 2004.

Basel III. Basel III counterparty credit risk and exposures to central counterparties - frequently asked questions. Basel Committee on Banking Supervision, Bank for International Settlements , 2012.

D. Bates. Post-'87 crash fears in the S&P 500 futures option market. *Journal of Econometrics* 94, 181–238, 2000.

G. Bekaert and E. Engstrom. Asset return dynamics under bad environment good environment fundamentals. Working paper, 2009.

L. Benzoni, P. Collin-Dufresne, and R. Goldstein. Explaining asset pricing puzzles associated with the 1987 market crash. *Journal of Financial Economics* 101, 552–573, 2011.

P. Bernstein. *The Power of Gold: The History of an Obsession* (Wiley), 2012.

D. Bertsimas, L. Kogan, and A. Lo. When is time continuous? *Journal of Financial Economics* 55, 173–204, 2000.

H. Bhamra, L.-A. Kuhn, and I. Strebulaev. The levered equity risk premium and credit spreads: A unified framework. *Review of Financial Studies* 23, 645–703, 2010.

F. Black and M. Scholes. The pricing of options and corporate liabilities. *Journal of Political Economy* 81, 637–654, 1973a.

F. Black and M. Scholes. The pricing of options and corporate liabilities. *Journal of Political Economy* 81, 637–654, 1973b.

W. Black. *The Platinum Group Metals Industry* (Woodhead Publishing Limited), 2000.

R. R. Bliss and N. Panigirtzoglou. Option-implied risk aversion estimates. *Journal of Finance* 59, 407–446, 2004.

N. Bloom. The impact of uncertainty shocks. *Econometrica* 77, 623–685, 2009.

T. Bollerslev, G. Tauchen, and H. Zhou. Expected stock returns and variance risk premia. *Review of Financial Studies* 22, 4463–4492, 2009.

J. Boudoukh, R. Michaely, M. Richardson, and M. Roberts. On the importance of measuring payout yield: Implications for empirical asset pricing. *Journal of Finance* 62, 877–915, 2007.

N. Branger, A. Kraftschik, and C. Völkert. The fine structure of variance: Consistent pricing of VIX derivatives. Working paper, 2014.

D. Breeden and R. Litzenberger. Prices of state contingent claims implicit in options prices. *Journal of Business* 51, 621–651, 1978.

M. Brennan. The supply of storage. *American Economic Review* 48, 50–72, 1958.

M. Britten-Jones and A. Neuberger. Option prices, implied price processes, and stochastic volatility. *Journal of Finance* 51, 621–651, 2000.

M. Broadie, M. Chernov, and M. Johannes. Model specification and risk premia: Evidence from futures options. *Journal of Finance* 62, 1453–1490, 2007.

J. Campbell. Stock returns and the term structure. *Journal of Financial Economics* 18, 373–399, 1987.

J. Campbell, 1999. Asset prices, consumption and the business cycle. in John B. Taylor and Michael Woodford, (eds.) *Handbook of Macroeconomics*. volume 1 (Elsevier Science, North Holland, Amsterdam).

J. Campbell, S. Giglio, C. Polk, and R. Turley. An intertemporal capm with stochastic volatility. Working paper, 2012.

J. Campbell and R. Shiller. Stock prices, earnings, and expected dividends. *Journal of Finance* 43, 661–676, 1988.

J. Campbell and S. Thompson. Predicting excess stock returns out of sample: Can anything beat the historical average? *Review of Financial Studies* 21, 1509–1531, 2008.

L. Canina and S. Figlewski. The informational content of implied volatility. *Review of Financial Studies* 6, 659–681, 1993.

J. Cao and B. Han. Cross-section of option returns and idiosyncratic stock volatility. *Journal of Financial Economics* 20, 1–10, 2013.

P. Carr and D. Madan. Towards a theory of volatility trading. in *Option Pricing, Interest Rates, and Risk Management*. 417–427, 1998.

P. Carr and D. Madan. Option pricing and the fast fourier transform. *Journal of Computational Finance* 2, 61–73, 1999.

P. Carr and J. Yu. Risk, return, and Ross recovery. *Journal of Derivatives* 20, 2012.

J. Casassus and P. Collin-Dufresne. Stochastic convenience yield implied from commodity futures and interest rates. *Journal of Finance* 60, 2283–2331, 2005.

G. Cawthorn. Seventy-fifth anniversary of the discovery of the platiniferous merensky reef. *Platinum Metals Review* 43, 146–148, 1999.

B. Y. Chang, P. Christoffersen, and K. Jacobs. Market skewness risk and the cross-section of stock returns. *Journal of Financial Economics* 107, 46–68, 2013.

M. Chernov, R. Gallant, E. Ghysels, and G. Tauchen. Alternative models for stock price dynamics. *Journal of Econometrics* 116, 225–257, 2003.

B. Christensen and N. Prabhala. The relation between implied and realized volatility. *Journal of Financial Economics* 50, 125–150, 1998.

P. Christoffersen, K. Jacobs, and K. Mimouni. Volatility dynamics for the S&P 500: Evidence from realized volatility, daily returns, and option prices. *Review of Financial Studies* 23, 3141–3189, 2010.

T. Clark and K. West. Approximately normal tests for equal predictive accuracy in nested models. *Journal of Econometrics* 138, 291–311, 2007.

J. Cochrane. The dog that did not bark: A defense of return predictability. *Review of Financial Studies* 21, 1533–1575, 2008.

J. Coval and T. Shumway. Expected option returns. *Journal of Finance* 56, 983–1009, 2001.

J. Cox, J. Ingersoll, and S. Ross. A theory of the term structure of interest rates. *Econometrica* 53, 385–407, 1985.

D. Cuoco and H. Liu. Optimal consumption of a divisible durable good. *Journal of Economic Dynamics and Control* 24, 561–613, 2000.

D. Dickey and W. Fuller. Distribution of the estimators for autoregressive time series with a unit root. *Journal of the American Statistical Association* 74, 427–431, 1979.

I. Drechsler. Uncertainty, time-varying fear, and asset prices. *Journal of Finance* 68, 1843–1889, 2013.

I. Drechsler and A. Yaron. What's vol got to do with it? *Review of Financial Studies* 24, 1–45, 2011.

S. Dubynskiy and R. S. Goldstein. Recovering drifts and preference parameters from financial derivatives. Working paper, 2013.

D. Duffie and C. Skiadas. Continuous-time asset pricing: A utility gradient approach. *Journal of Mathematical Economics* 23, 107–132, 1994.

D. Duffie and L. Epstein. Asset pricing with stochastic differential utility. *Review of Financial Studies* 5, 411–436, 1992.

D. Duffie, J. Pan, and K. Singleton. Transform analysis and asset pricing for affine jump-diffusions. *Econometrica* 68, 1343–1376, 2000.

P. Dybvig and S. Ross. Arbitrage. in M. Milgate J. Eatwell and P. Newman, (eds.) *The New Palgrave Dictionary of Economics* (Palgrave Macmillan), 1987.

G. Elliot and J. Stock. Inference in time series regression when the order of integration of a regressor is unknown. *Econometric Theory* 10, 672–700, 1994.

R. Engle and C. Granger. Co-integration and error correction: Representation, estimation, and testing. *Econometrica* 55, 251–276, 1987.

L. Epstein and S. Zin. Substitution, risk aversion, and the temporal behavior of consumption and asset returns: A theoretical framework. *Econometrica* 57, 937–969, 1989.

B. Eraker. Do stock prices and volatility jump? Reconciling evidence from spot and option prices. *Journal of Finance* 59, 1367–1403, 2004.

B. Eraker, M. Johannes, and N. Polson. The impact of jumps in returns and volatility. *Journal of Finance* 53, 1269–1300, 2003.

B. Eraker and I. Shaliastovich. An equilibrium guide to designing affine pricing models. *Mathematical Finance* 18, 519–543, 2008.

C. Erb and C. Harvey. The golden dilemma. *Financial Analysts Journal* 69, 10–42, 2013.

E. Fama and K. French. Business cycles and the behavior of metals prices 43, 1075–1093, 1988.

E. Fama and K. French. Business conditions and expected returns on stocks and bonds. *Journal of Financial Economics* 25, 1989.

E. Fama and J. MacBeth. Risk, return, and equilibrium: Empirical tests. *Journal of Political Economy* 81, 607–636, 1973.

J. Fernandez-Villaverde and J. F. Rubio-Ramírez. Macroeconomics and volatility: Data, models, and estimation. in *Advances in Economics and Econometrics: Theory and Applications,* (Cambridge University Press), 2013.

S. Figlewski. Estimating the implied risk neutral density for the u.s. market portfolio. in Jeffrey R. Russell Timothy Bollerslev and Mark Watson, (eds.) *Volatility and Time Series Econometrics: Essays in Honor of Robert F. Engle* (Oxford University Press), 2008.

A. Frazzini and L. Pedersen. Embedded leverage. Working paper, 2012.

X. Gabaix. Variable rare disasters: An exactly solved framework for ten puzzles in macro-finance. *Quarterly Journal of Economics* 127, 645–700, 2012.

J. Gomes, L. Kogan, and M. Yogo. Durability of output and expected stock returns. *Journal of Political Economy* 117, 941–986, 2009.

F. Gourio. Disaster risk and business cycles. *American Economic Review* 102, 2734–2766, 2012.

A. Goyal and I. Welch. A comprehensive look at the empirical performance of equity premium prediction. *Review of Financial Studies* 21, 1455–1508, 2008.

F. Guvenen. Mismeasurement of the elasticity of intertemporal substitution: The role of limited stock market participation. Unpublished manuscript, University of Rochester, 2001.

R. Hall. Intertemporal substitution in consumption. *Journal of Political Economy* 96, 339–357, 1988.

L. P. Hansen, J. Heaton, J. Lee, and N. Roussanov. Chapter 61: Intertemporal substitution and risk aversion. in *Handbook of Econometrics, Volume 6, Part A*. 3967, 2007.

L. P. Hansen, J. C. Heaton, and N. Li. Consumption strikes back? Measuring long-run risk. *Journal of Political Economy* 116, 260–302, 2008.

L. P. Hansen and T. Sargent. Fragile beliefs and the price of model uncertainty. Working paper, 2006.

L. P. Hansen and J. Scheinkman. Long-term risk: An operator approach. *Econometrica* 77, 177–234, 2009.

L. P. Hansen and K. Singleton. Generalized instrumental variables estimation of nonlinear rational expectations model. *Econometrica* 50, 1269–1286, 1982.

H. Hasseltoft. Stocks, bonds and long-run consumption risks. *Journal of Financial and Quantitative Analysis* 47, 309–332, 2012.

S. Heston. A closed-form solution for options with stochastic volatility with applications to bonds and currency options. *Review of Financial Studies* 6, 327–343, 1993.

S. Heston, R. Korajczyk, and R. Sadka. Intraday patterns in the cross-section of stock returns. *Journal of Finance* 65, 1369–1407, 2010.

M. Hirschey and J. Nofsinger. *Investments: Analysis and Behavior* (McGraw-Hill), 2008.

R. Hodrick. Dividend yields and expected stock returns: Alternative procedures for inference and measurement. *Review of Financial Studies* 5, 357–386, 1992.

X. Huang and G. Tauchen. The relative contribution of jumps to total price variation. *Journal of Financial Econometrics* 3, 456–499, 2005.

J. Hull. *Options, Futures, and Other Derivatives* (Prentice Hall, 8th edition), 2011.

J. C. Jackwerth. Recovering risk aversion from option prices and realized returns. *Review of Financial Studies* 13, 433–451, 2000.

G. Jiang and Y. Tian. The model-free implied volatility and its information content. *Review of Financial Studies* 18, 1305–1342, 2005.

S. Johansen. Statistical analsyis of cointegration vectors. *Journal of Economic Dynamics and Control* 12, 231–254, 1988.

S. Johansen. *Likelihood-Based Inference in Cointegrated Vector Autoregressive Models* (Oxford: Oxford University Press), 1995.

J. Jurek. Crash-neutral currency carry trades. *Journal of Financial Economics* 113, 325–347, 2014.

S. Kandel and R. Stambaugh. On the predictability of stock returns: An asset allocation perspective. *Journal of Finance* 51, 385–424, 1996.

D. Keim and R. Stambaugh. Predicting returns in the stock and bond markets. *Journal of Financial Economics* 17, 357–390, 1986.

B. Kelly and H. Jiang. Tail risk and asset prices. *Review of Financial Studies* 27, 2841–2871, 2014.

D. M. Kreps and E. L. Porteus. Temporal resolution of uncertainty and dynamic choice theory. *Econometrica* 46, 185–200, 1978.

V. Lannoye. *The History of Money for Understanding Economics* (CreateSpace Independent Publishing), 2011.

LBMA. Gold lease rates: A rapidly changing market. *Alchemist* 53, 2009.

LBMA and LPPM. A guide to the london precious metals markets , 2008.

A. Le and H. Zhu. Risk premia in gold lease rates. Working paper, 2013.

M. Lettau and S. Ludvigson. Consumption, aggregate wealth, and expected stock returns. *Journal of Finance* 56, 815–849, 2001.

A. Lewis. *Option Valuation Under Stochastic Volatility* (Finance Press), 2000.

Y. Li, D. Ng, and B. Swaminathan. Predicting market returns using aggregate implied cost of capital. *Journal of Financial Economics* 110, 419–436, 2013.

J. Lintner. Inflation and security returns. *Journal of Finance* 30, 259–280, 1975.

J. Liu, J. Pan, and T. Wang. An equilibrium model of rare-event premia and its implication for option smirks. *Review of Financial Studies* 18, 131–164, 2005.

L. Liu, A. Patton, and K. Sheppard. Does anything beat 5-minute RV? A comparison of realized measures across multiple asset classes. Working paper, 2013.

Z. Lu and Y. Zhu. Volatility components: The term structure dynamics of VIX futures. *Journal of Futures Markets* 30, 230–256, 2010.

R. Lucas. Asset prices in an exchange economy. *Econometrica* 46, 1429–1445, 1978.

H. Lustig, S. V. Nieuwerburgh, and A. Verdelhan. The wealth-consumption ratio. Working paper, 2012.

A. Madison. *The World Economy: Historical Statistics* (Paris, OECD), 2003.

A. Manela and A. Moreira. News implied volatility and disaster concerns. Working paper, 2014.

R. Mehra and E. Prescott. The equity premium: A puzzle. *Journal of Monetary Economics* 10, 211–219, 1985.

G. Milstein. *Numerical Integration of Stochastic Differential Equations* (Kluwer Academic, Boston), 1995.

E. Nakamura, J. Steinsson, R. Barro, and J. Ursua. Crises and recoveries in an empirical model of consumption disasters. *American Economic Journal: Macroeconomics* 5, 35–74, 2013.

W. Newey and K. West. A simple, positive semi-definite, heteroskedasticity and autocorrelation consistent covariance matrix. *Econometrica* 59, 817–858, 1987.

S. Ng and P. Perron. Unit root tests in arma models with data-dependent methods fort the selection of the truncation lag. *Journal of the American Statistical Association* 90, 268–281, 1995.

M. Nowotny. Disaster begets crisis: The role of contagion in financial markets. Working paper, 2011.

M. Ogaki and C. Reinhart. Measure intertemporal substitution: The role of durable goods. *Journal of Political Economy* 106, 1078–1098, 1998.

J. Pan. The jump-risk premia implicit in options: Evidence from an integrated time-series study. *Journal of Financial Economics* 63, 3–50, 2002.

L. Pastor, M. Sinha, and B. Swaminathan. Estimating the intertemporal risk-return tradeoff using the implied cost of capital. *Journal of Finance* 63, 2859–2897, 2008.

L. Pastor and R. Stambaugh. Liquidity risk and expected stock returns. *Journal of Political Economy* 111, 642–685, 2003.

T. Rietz. The equity risk premium: A solution. *Journal of Monetary Economics* 22, 117–131, 1988.

S. Ross. The recovery theorem. forthcoming in Journal of Finance, 2013.

M. Rubinstein. Implied binomial trees. *Journal of Finance* 49, 771–818, 1994.

E. Schwartz. The stochastic behavior of commodity prices: Implications for valuation and hedging. *Journal of Finance* 52, 923–973, 1997.

S. B. Seo and J. Wachter. Option prices in a model with stochastic disaster risk. Working paper, 2014.

I. Shaliastovich. Learning, confidence and option prices. Working Paper, 2010.

J. Shanken. On the estimation of beta-pricing models. *Review of Financial Studies* 5, 1–55, 1992.

D. Shimko. Bounds of probability. *RISK* 6, 33–37, 1993.

Z. Song. Expected VIX option returns. Working paper, 2013.

Z. Song and D. Xiu. A tale of two option markets: Pricing kernels and volatility risk. Working paper, 2013.

R. F. Stambaugh. Predictive regressions. *Journal of Financial Economics* 54, 375–421, 1999.

J. Stein. Overreactions in the options market. *Journal of Finance* 44, 1011–1023, 1989.

J. Stock and M. Watson. A simple estimator of cointegrating vectors in higher order integrated systems. *Econometrica* 61, 783–820, 1993.

G. Tauchen and R. Hussey. Quadrature-based methods for obtaining approximate solutions to nonlinear asset pricing models. *Econometrica* 59, 317–396, 1991.

G. Tauchen and V. Todorov. Volatility jumps. *Journal of Business and Economic Statistics* 29, 235–371, 2011.

P. Thomas and E. Boyle. Gold availability - world, a minerals availability program appraisal. *U.S. Bureau of Mines, Washington D.C.* , 1986.

W. Torous, R. Valkanov, and S. Yan. On predicting stock returns with nearly integrated explanatory variables. *Journal of Business* 77, 937–966, 2004.

P. Tufano. Who manages risk? An empirical examination of risk management practices in the gold mining industry. *Journal of Finance* 51, 1097–1137, 1996.

G. Vilkov and Y. Xiao. Option-implied information and predictability of extreme returns. Working paper, 2012.

A. Vissing-Jørgensen. Limited asset market participation and the elasticity of intertemporal substitution. *Journal of Political Economy* 110, 825–853, 2002.

A. Vissing-Jørgensen and O. Attanasio. Stock market participation, intertemporal substitution and risk aversion. *American Economic Review (Papers and Proceedings)* , 2003.

J. Wachter. Can time-varying risk of rare disasters explain aggregate stock market volatility? *Journal of Finance* 68, 987–1035, 2013.

P. Weil. The equity premium puzzle and the risk-free rate puzzle. *Journal of Monetary Economics* 24, 401–421, 1989.

J. Wright and H. Zhou. Bond risk premia and realized jump risk. *Journal of Banking and Finance* 33, 2036–2049, 2009.

M. Yogo. A consumption-based explanation of expected stock returns. *Journal of Finance* 61, 539–580, 2006.

J. Zhang and Y. Zhu. VIX futures. *Journal of Futures Markets* 26, 521–531, 2006.